NASA Contractor Report 178213

DESIGN and EVALUATION of a FAILURE DETECTION and ISOLATION ALGORITHM for RESTRUCTURABLE CONTROL SYSTEMS

Jerold L. Weiss, John Y. Hsu

ALPHATECH, Inc.
Burlington, Massachusetts

March, 1987

Prepared For
NASA Langley Research Center
Under Contract NAS1-18004

National Aeronautics and
Space Administration

Langley Research Center
Hampton, Virginia 23665

CONTENTS

CONTENTS (Continued)

CONTENTS (Continued)

CONTENTS (Continued)

FIGURES

FIGURES (Continued)

FIGURES (Continued)

FIGURES (Continued)

FIGURES (Continued)

TABLES

SECTION 1

INTRODUCTION

This report presents the result of a one-year effort sponsored by the NASA Langley Research Center under contract NAS1-18004 to design and evaluate a Failure Detection and Isolation (FDI) algorithm for application to restructurable flight control. The restructurable or reconfigurable flight control system (RFCS) concept is a fault tolerant control concept which is capable of automatically generating the control action needed for recovery from unanticipated emergencies as well as providing the stability and control augmentation for controllable flight under these circumstances. Under NASA sponsorship, ALPHATECH, Inc., has been developing and testing many component technologies which will be necessary for near term demonstration and operational development of the RFCS concept. Currently, the FDI system developed for this contract is being integrated with other RFCS components for demonstration on NASA's modified B-737 simulation, (see [1] - [4] for a general discussion of the RFCS concept).

The FDI function is a critical component to the RFCS concept because there will always be conditions which can not be handled by a normal (i.e., any acceptable normal) control system, (which, of course, includes the pilot). Those failures which can not be so handled must be detected so that the RFCS knows when to reconfigure,and these failures must be isolated or identified so that the proper reconfiguration action is taken. Although there are a variety

of failures which may be important, those which result in a loss of control authority are most important since they not only result in emergency conditions, but also impact how well one can reliably respond to such conditions. This effort concentrated on the development of an FDI system to handle all such failures. These failures are known generically as control element failures and include (though are not limited to) runaway, stuck, floating, and partially missing surfaces, as well as engine failures such as loss of thrust and stuck throttle failure modes.

This project was divided into four major tasks. Task 1 addressed the issue of fundamental limits to FDI performance for the decentralized FDI approach previously developed by ALPHATECH, Inc. Task 2 was a preliminary design effort which utilized ALPHATECH's design methodology and a preliminary assessment of errors. Task 3 was an algorithm refinement phase which utilized simulation results to provide better estimates of errors for use in the design methodology, and Task 4 was an assessment of critical issues for further development. The availability of flight recorded data from flights of NASA's transportation systems research vehicle (TSRV) and NASA's modified simulation for that aircraft motivated our application to the B-737 aircraft. The flight recorded data also provided a unique opportunity for realistic evaluation of performance limitations in Task 1.

1.1 CONTRIBUTIONS

The detection and isolation of generic control element failures has received significant attention (e.g., see [5] - [13]) not only due to its importance for RFCS but also because of the difficulties associated with the need to use analytical redundancy in the solution method. Analytical

redundancy (unlike direct or hardware redundancy) refers to the concept of comparing dissimilar sensors using analytical or mathematical relationships between those sensors. Analytical redundancy is necessary for many control element failure modes because direct redundancy or built-in test equipment (BITE) is not available (consider a partially missing surface, for example). Furthermore, those control element failures which might also be amenable to direct redundancy and/or BITE (e.g., a stuck surface due to a loss of hydraulic pressure) can sometimes be handled more efficiently using analytical redundancy. This is because BITE, by definition, tests only the "preconditions" which are necessary for system operation (e.g., power applied). Analytical redundancy, however, tests the functionality of a particular system, thereby encompassing all modes of failures. Furthermore, it does this without hardware duplication, thereby reducing initial cost and weight and increasing overall system reliability and maintainability (fewer pieces of hardware to fail) as well.

Unfortunately, the design of FDI systems using analytical redundancy is difficult because of the sometimes significant inaccuracies associated with the mathematical models which are employed. This fact provided the motivation for ALPHATECH's development of an FDI design methodology which addresses the impact of such errors. One major contribution of this effort was the refinement and application of this design methodology for the control element FDI problem. This methodology consists of methods for choosing the structure of an FDI algorithm, optimizing its parameters in the presence of unavoidable modeling errors, and performing sensitivity analyses. It is largely based on the notion of discrimination metrics which can be used to bound, on an average basis, the decision errors associated with an FDI process. Such analyses

require a statistical model of system behavior which includes a wide variety of modeling errors (including parametric errors, unmodeled dynamics, nonlinearities, etc.). A qualitative analysis and several "error budgets" are used to derive these descriptions.

The other major contribution is the development and demonstration of an FDI system for detecting and isolating all important control element failure modes. This system is an advance over systems which only handle a limited class of failure modes (e.g., [51]). The key element in developing such a system is the recognition that the failures which are important are those which result in "large" failure signatures (as later defined in this report) but that different failure modes give rise to different temporal signature characteristics which are unknown a priori. The FDI system developed for this project uses only qualitative information about failure signatures (we assume they are coherent, although this is not a requirement for the design) and detects and isolates failures using only failure "size" information.

The general FDI concept used in this project is known as a decentralized approach because of our attempt to assess system redundancy and utilize only the most well-known parts of the system for specific FDI tasks. This is in contrast to centralized methods which integrate all information in an "optimal" manner. Such methods work well in ideal circumstances but frequently run into trouble when model error exists. The loss of optimality under ideal circumstances, which results from using the decentralized approach, is more than compensated by the increase in robustness to modeling errors. The first decentralization employed in this project is the decomposition of the control element FDI problem into two separate subproblems; actuator-path FDI and aircraft-path FDI. The actuator path problem is a "local" FDI problem which

is concerned only with failures which occur between the location of (total) actuator command measurements (e.g., in the flight control computer) and the location of actuator output measurements (e.g., on a control rod). The aircraft path problem is concerned with all failures which occur outboard of the actuator output measurement. Two decoupled subsystems were designed to handle each of these subproblems (see Section 2). The actuator path subsystem consists of decoupled actuator path systems (one for each actuator) which make use of actuation models. It can handle any combination of sequential or simultaneous actuator path failures. The aircraft path subsystem utilizes models which relate aircraft motion to the measured control values and can only handle single failures. The ability to handle multiple aircraft path failures depends on knowledge of failure signatures which is not available without explicit control excitation (e.g., dither signals). Such an approach was not considered in this work.

1.2 OUTLINE OF THIS REPORT

Section 2 formulates the overall control element FDI problem in terms of the failure modes of interest, the goals and assumptions used for FDI design and the various configuration options available for handling the actuator-path and aircraft-path subproblems. Section 3 describes the decentralized approach to FDI and includes a variety of examples of how decentralized "residuals" can be generated and provides a decision structure which takes maximal advantage of these residuals. A design methodology for this structure is presented and many examples of hypothesis test designs which will be useful for the control element FDI problem are given. Section 4 presents a method for evaluating fundamental limits to FDI performance using discrimination metrics and solves

an example which is used in Section 5. Section 5 applies the concepts developed in Sections 3 and 4 to the B-737 aircraft and goes through design and analysis details and simulation results for both actuator-path and aircraft-path subsystems. Conclusions and recomendations are given in Section 6.

SECTION 2

PROBLEM FORMULATION

The broad goal of this project is to develop a system for detecting the occurrence of any control element failure (to be defined) and isolating or identifying the affected control. Major failures should be detected quickly so that appropriate reconfiguration can take place, but false alarms must also be minimized.

We will assume single flight condition operation, full measurement of the rigid body state vector (e.g., measurements of body referenced angular rates and relative wind, as well as body referenced accelerometer measurements), and models (not necessarily linear models) of the aircraft and actuator. Sensor errors including noise, scale factor and bias (within some design specification) and model errors including parameter errors and unmodeled dynamics (also within some design spec.) must be accounted for in the design process.

The motion of the aircraft in response to control action under <u>any</u> operational status (failed or unfailed) is now assumed to be representable by a set of differential equations, viz.

$$\dot{x} = f(x, \delta_E) \tag{2-1a}$$

$$y = h(x, \delta_E) \tag{2-1b}$$

where x is some n-dimensional state vector that includes the effects of flexible modes and disturbances due to turbulence, in addition to the rigid body

states which describe body referenced motion of the aircraft through the atmosphere; δ_E is a vector of "effective" control values; and y is the vector of measurable quantities. Equation 2-1 is true independent of the aircraft's operational status.

Control element failures are now construed to mean: δ_E does not "follow" the desired control commands. That is, each of the, m, effective controls, $\delta_E{}^i$, is derived from an independent set of differential equations which do depend on the status of the aircraft (failed or not failed). These equations written in operator form are

$$\delta_E{}^i = g_E{}^i\{\delta_A{}^i\} + d_E{}^i \qquad (2\text{-}2)$$

$$\delta_A{}^i = g_A{}^i\{\delta_c{}^i\} + d_A{}^i \qquad (2\text{-}3)$$

where δ_c represents the control commands being input to the actuation mechanism, δ_A represents the output of the actuation mechanism, $g_A{}^i$ and $g_E{}^i$ are causal operators, and d_E and d_A are time varying "disturbance" functions. under no-failure conditions Eqs. 2-2 and 2-3 reduce to,

$$\delta_E{}^i = \delta_A{}^i \qquad (2\text{-}4)$$

$$\delta_A{}^i = g_{OA}{}^i\{\delta_c{}^i\} \qquad (2\text{-}5)$$

for all i, where $g_{OA}{}^i$ is a model of a working actuation mechanism.*

*The definition of each actuator's input and output (δ_c and δ_A) is not unique. For control surfaces, inputs include DFCS outputs, electrical signals between a DFCS and an actuator, and differential pressure in a hydraulic actuator. Outputs could be taken at the actuator output, on a control rod, or at the surface hinge. For this project we have assumed that measurements of δ_A and δ_c are available. Therefore, we can define each actuator's input and output by the location of the corresponding measurements.

Equations 2-2 and 2-3 are sufficient for describing all types of control element failures. These failures are decomposed into two categories. "Actuator-path" failures are those failures in which Eq. 2-3 differs from Eq. 2-5 and "aircraft-path" failures are those in which Eq. 2-2 is different from Eq. 2-4. Tables 2-1 and 2-2 describe a variety of control element failure models in terms of Eqs. 2-2 and 2-3. Failure mechanisms which result in behavior characterized by each of the models in Table 2-1 can be conceived. In general, models and mechanisms depend on the exact locations of actuator input and output measurements.

TABLE 2-1. ACTUATOR PATH FAILURES

Stuck	$g_A(\delta) = 0$	d_A = Constant
Floating	$g_A(\delta) = 0$	d_A = Follows local α
Runaway	$g_A(\delta) = 0$	d_A = Slews to limit
Reduced Bandwidth	$g_A(\delta) = g_{AF}(\delta)$	$d_A = 0$

TABLE 2-2. AIRCRAFT PATH FAILURES

Stuck	$g_E(\delta) = 0$	d_E = Constant
Floating	$g_E(\delta) = 0$	d_E = Follows local α
Runaway	$g_E(\delta) = 0$	d_E = Slews to limit
Partial Loss (A)	$g_E(\delta) = k \cdot \delta$	$d_E = 0$
(B)	$g_E(\delta) = k \cdot \delta$	d_E = Follows local α

($k = 1$ - fraction-lossed)

The availability of actuation input and output measurements has an impact on the configuration of decentralized FDI systems. A centralized system would simply combine the actuation models with the aircraft state equation in Eq. 2-1 and add the output measurements to the observation vector. In the decentralized approach, an assessment of the "redundancy" available from each part of the model is made and possible system decompositions (which utilize only subsets of models) are explored. Actuation output measurements allow consideration of separate actuator and aircraft-path subsystems as described below.

Figure 2-1 describes the information flow which is available for FDI for some measurement configuration. Several parallel actuator paths are shown. Failures in each actuator can be independently detected through the use of the analytical redundancy which is embedded in the independent actuator models. That is, actuator-path failures can be detected by comparing a predicted actuator output (based on the measured input and an actuator model) with the measured output.

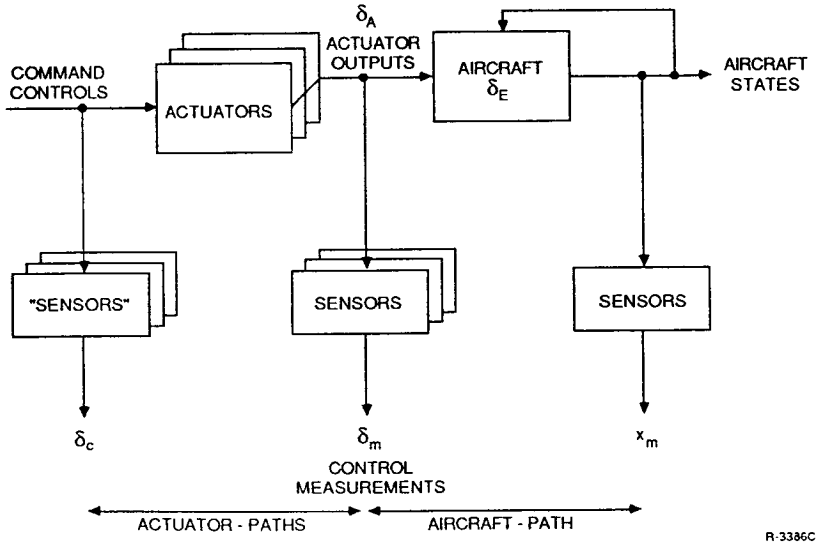


Fig. 2-1. Measurement Configuration and Analytic Redundancy Implications

When an effective control value (i.e., the control value which actually moves the airplane) differs from the measured output of the actuator, then an aircraft-path failure exists. These failures can be detected by the use of the analytical redundancy which is embedded in an aircraft model. That is, aircraft-path failures can be detected by comparing the measured motion variables (which are a function of the aircraft states) with a prediction of these variables based on the control measurements.

Clearly, from the figure, all control element failures (both actuator-path and aircraft-path) could be detected using an aircraft model that includes the actuator models thereby eliminating the need for actuator output measurements and reducing the cost and weight associated with the sensor hardware and redundancy management. Furthermore, such a system results in an FDI algorithm which is based on sensors which are more likely to survive potential causes of failures (e.g., battle damage, sabotage, etc.).

On the other hand, the parallel actuator path FDI algorithms tend to be very simple and more reliable than the aircraft path algorithms. Also, on most commercial and military aircraft, the cost and weight of establishing actuator output measurements of some kind would not be prohibitive; (in fact, any servo driven actuator already has an output measurement, although not necessarily the best one in terms of failure mode coverage). As a result, we expect that independent actuator path algorithms will be an important part of control element FDI and have, therefore, considered their development for this project.

Given the development of an actuator-path FDI system, the next question is how to cover the remaining aircraft-path failure modes. There are two options in this regard. The first option is to create a backup system to the

actuator-path algorithm by incorporating the actuator models in the aircraft model and using control commands as inputs. This would then cover both aircraft-path and actuator-path failures. The second option is to utilize the control measurements for prediction of aircraft motion in the aircraft path algorithm. Since the latter option requires no ambiguity resolution (aircraft-path and actuator-path algorithms cannot disagree) and since this option may provide a more reliable aircraft-path algorithm (no actuator model errors), we have developed this second option for this project.

Figure 2-2 shows the resulting high-level structure for the FDI algorithm being developed. At the top of this figure, actuator models and measurements are used to compare predicted and measured actuator outputs. This comparison consists of independent deflection residual signals. When the i^{th} deflection residual is large, a failure of the i^{th} control element is indicated. The decision processes for actuator path failures are decoupled and are responsible for deciding if each residual signal is large because of a failure or because of model error excitation. In the lower part of the figure, measurements of various aircraft states, along with the measured deflections and an aircraft model, are used to form residual signals which correspond to the six forces and moments which define the aircraft motion (details given in Section 5). The aircraft path decision process is then responsible for detecting when these residuals are larger than expected (accounting for noise and model error) and for deciding which control element failure is responsible for the force and moment imbalance indicated by the relative sizes of the six residuals.

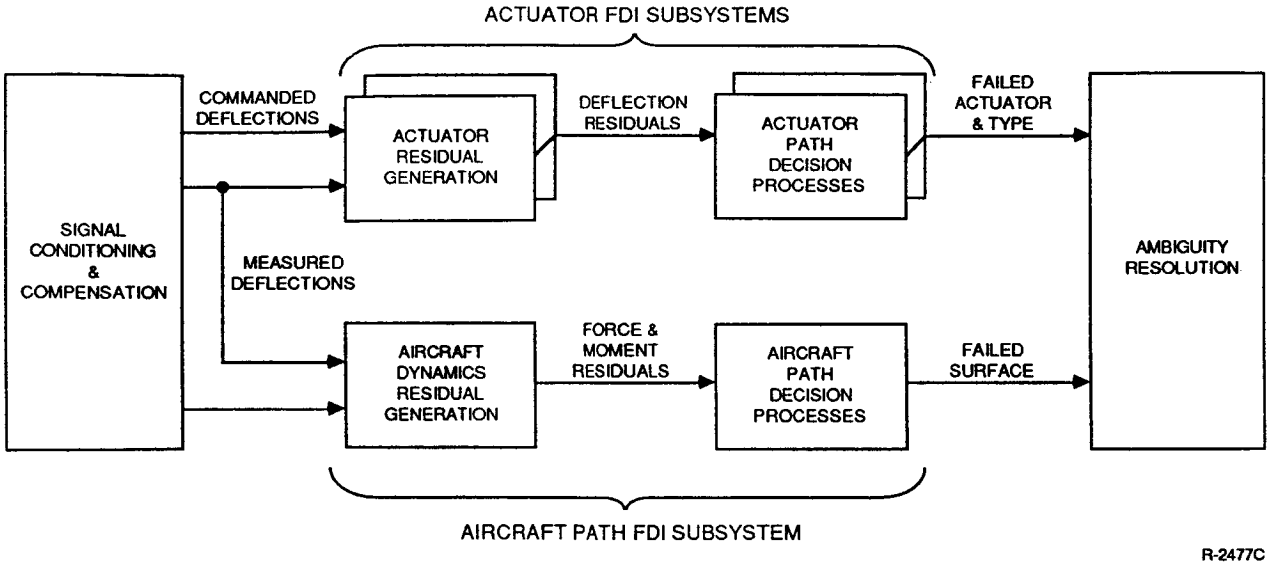


Figure 2-2. FDI Structure

SECTION 3

A DECENTRALIZED APPROACH TO FDI

In this section we motivate the need for a decentralized approach to the failure detection and isolation (FDI) problem. The term "decentralized" as applied to the FDI problem is used to indicate the uncoupled use of parts of a system model to develop "redundancy relationships," and the selective use of these relations in the FDI process. This decentralized approach was first used in the NASA F-8 sensor FDI project, [14] - [16], in order to allow clear trade-offs between model error and failure sensitivity to be made. Such trade-offs are the key to successful FDI design for systems which cannot be precisely described by the kinds of models which form the basis of the many "optimal" FDI methods, (e.g., see [17]). This is because modeling error always exists (including parametric errors, unmodelled dynamics, and nonstationary inputs) and because the optimal methods, which guarantee optimality under ideal conditions, are typically not "robust."

The success of the F-8 work spawned many research activities into robust FDI methods. Willsky and his co-workers at MIT, ([18] - [21] first addressed the residual generation problem in terms of solving for "parity checks" (which can be interpreted as auto-regressive-moving-average models) which are insensitive to model errors. Pattipati and co-workers at ALPHATECH ([22], [24]) extended these ideas to include the trade-off between model error and detection sensitivity and provided a unified framework for developing robust FDI methods. Weiss and co-workers at ALPHATECH, at the same time, developed a

control element FDI algorithm based on the conceptual framework of the F-8 sensor work, [5]. This algorithm provided a generic structure for robust decisionmaking and, more importantly, began to address the need for robustness optimization and sensitivity analysis in the design of decision processes. These ideas have been extended during this project and are detailed in subsection 3.3.

3.1 OVERVIEW

- What is failure detection and isolation (FDI)?

Obviously, FDI deals with the problem of detecting deviations from normal behavior in specified components (sensors or effectors) and isolating the particular component which has "failed." The key point in this sentence is that in order to detect and isolate deviations, one requires a specification or model of "normal behavior" and of the anomalous behavior to be detected. Furthermore, for each type of anomaly to be detected, these models must provide sufficient redundant information to allow one to detect each anomaly and to distinguish it from others. For example, in a triplex sensor system, in which there are three identical sensors of each type, one can perform voting by examining each triple to determine if its components are consistent (i.e., normal). If one sensor in the triple is significantly different from the other two, then we could conclude that it has failed. In this scheme, the model information used is that the three sensors measure the identical quantity, and the model of a deviation can be specified in several ways, such as in terms of manufacturer's instrument specifications. As a second example, consider a relatively simple and often-used check in which successive samples of the output of a particular sensor are examined to determine if there is an

obvious inconsistency. Here the model information used is a crude measure of the bandwidth of the variable being sensed. Finally, consider a simple system involving linear motion and in which one has a velocity sensor and an accelerometer. Here the kinematic model v = a provides a mechanism for obtaining one redundant relationship between these sensors.

In the terminology used by Chow and Willsky [18], [20], the three examples just described are illustrations of direct (or hardware) redundancy, temporal (or self-test) redundancy, and analytic (or functional) redundancy, respectively. While there are clear differences among them, it is their similarities--in terms of being based on models and, more explicitly, on redundancy imbedded in those models--that we wish to stress. This permits us to construct a unified framework in which to examine and compose different approaches to failure detection and their robustness properties.

- What does an FDI algorithm do?

Roughly speaking, all failure detection systems can be described in terms of the conceptual block diagram of Fig. 3-1. This diagram has been used quite often (e.g., see [5], [23], [24], [18], [20]) and provides a framework for the design and analysis of robust FDI systems. In Fig. 3-1, there are two fundamental parts of failure detection. The first of these is the generation of sets of signals (called residuals) whose deviation from "normal behavior" (typically meaning near zero without significant trends or patterns during no failure operation) can be used as the basis for detecting and identifying system failures. The second component of a failure detection system is the decision process consisting of information collection and decision-logic functions. Here, the residuals that have been generated are processed in order to make FDI decisions.

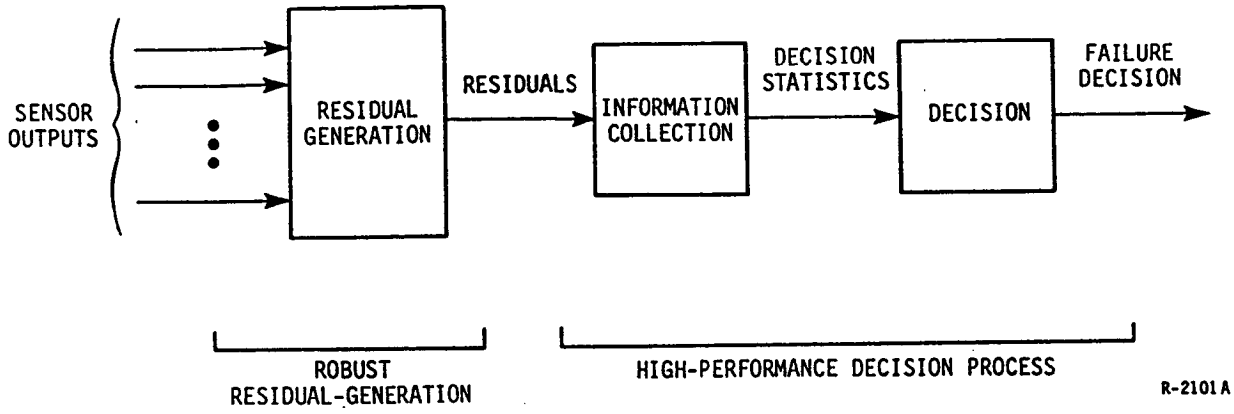


Figure 3-1. General Structure for FDI

A number of issues arises in the design of each of these subsystems, and to begin our discussion, let us focus first on the residual generation process.

3.1.1 Residual Generation

The way in which residuals can be generated varies greatly. For example, in a triplex system, if $y_1(k)$, $y_2(k)$, and $y_3(k)$ denote the outputs of three identical sensors, then $r_1(k) = y_1(k)-y_2(k)$ and $r_2(k) = y_2(k)-y_3(k)$ can be thought of as the residuals used in a voting system. In other FDI methods which have been suggested (e.g., see surveys [17], [25]), Kalman filters may be used to generate the residuals. In some of these methods, such as the detection filter approach [21] and [26], [27] Kalman-like filters are designed, but with gains chosen in particular ways so as to make particular failures more readily apparent. The decentralized approach to FDI, which is the topic of this report, provides a generalization of the residual generation mechanism employed in the voting scheme described above. In this approach,

each piece of the system model is examined to determine individual relationships among the measured quantities. Residuals can then be generated from these individual relations.

Now, although the use of Kalman filters for generating residuals may allow us to bypass the explicit identification of system redundancy, the absence of a specific assessment of redundancy frequently creates many difficulties when Kalman filter approaches are applied in a "top-down" manner. One of the reasons for these difficulties is the issue of robustness or model-error tolerance. By their very nature, good failure detection algorithms attempt to generate signals that are sensitive to system anomalies (i.e., failures). Given that all residual generation mechanisms use models of the relationships among available (i.e., measured or commanded) signals, we immediately see that the possibility exists for these residuals to be sensitive to modeling errors as well as to system failures. What is needed, therefore, is an FDI method with selective sensitivity. That is, an FDI algorithm should be based on residuals which are maximally sensitive to failures and minimally sensitive to model errors. It is here that the Kalman filter approaches run into trouble. All residuals produced by a Kalman filter are produced using a centralized system model. By definition, therefore, they tend to mix together relationships that are known very well with those that are far more uncertain. For this reason centralized approaches that are optimal when models are well known become far from optimal when model uncertainty is taken into account.

As mentioned above, the term "decentralized," as applied to the generation of residuals, refers to the identification and extraction of each individual source of system redundancy which we call a redundancy or parity relation. Selective sensitivity is achieved by examining each such relationship to

determine its robustness (or lack thereof) to various possible model errors and its sensitivity to specific failure modes. The most reliable set of relations providing the desired coverage (i.e., capable of detecting and identifying a specified set of failures) can then be determined, and separate FDI tests can ben be designed. Since only the most reliable relationships are used in each test, the effects of model errors are minimized, and thus the FDI tests can be implemented reliably. Furthermore, since each redundancy relation typically involves only small subsets of the set of available signals and of the set of model parameters, the problem of failure misclassification is minimized, and the effects of particular worst-case model error scenarios can be easily analyzed in great detail. Finally, since the resulting FDI system consists of a collection of extremely simple, low-order sub-algorithms, the overall system becomes easier to implement, verify, troubleshoot (either for logical errors or to pinpoint weaknesses identified during optimal tests caused by unanticipated sources of error), and modify.

To summarize, in order to achieve selective sensitivity in an FDI algorithm, we generate residuals in a decentralized manner. Individual relationships between the measurable variables can then be considered in terms of their sensitivity to specific failure modes and to various sources of model errors. Only the "best" relationships for detecting and distinguishing individual failure modes or subsets of failure modes will then be used in the hypothesis tests which make up the second part of the FDI structure of Fig. 3-1, the decision process. Details of the various decentralized residual generation methods are given in subsection 3.2.

3.1.2 The Decision Process

The decision process accumulates information about the system operation, performs a variety of hypothesis tests, and combines the results of these tests into a logical decision about the system status.

The accumulation of information is necessitated by the fact that although the instantaneous value of a residual derived from any particular parity relation provides one piece of information about possible failures, typically this one piece of information is not sufficient for accurate detection and identification. Rather, the information contained in successive values of the residual must be accumulated over time in order to achieve acceptable levels of performance (see [20] for a discussion of the several ways in which information can be collected).

The fact that information must be accumulated over time, coupled with the fact that the failure onset time is unknown, creates a situation which has led to a variety of failure decision mechanisms. The reason that so many methods have been proposed stems from the considerable advantage obtained from knowing or estimating the failure onset time as described below.

UNKNOWN ONSET TIME

The advantage of knowing the failure onset time is easily illustrated using simple measures of failure distinguishability (see Section 4). Consider, for example, the distinguishability of a constant, nonzero, bias which occurs at an unknown time ("jump failure") and measurements which are contaminated by white Gaussian noise. One commonly used detection approach is to operate on a sliding window of data and declare a failure when the output of this operation exceeds a threshold.

In particular, let y(k) represent the observed signal that obeys

$$\begin{aligned} & y(k) = \eta(k) && \text{for all } k < k_f, \text{ and} \\ \text{under } H_0: \quad & y(k) = \eta(k) && \text{for all } k \geq k_f \\ \text{under } H_1: \quad & y(k) = m + \eta(k) && \text{for all } k \geq k_f \end{aligned} \tag{3-1}$$

Now, if the decision process is defined by applying a maximum likelihood hypothesis testing technique over the sliding window, a decision statistic, S_k (the log-likelihood ratio), is formed by,

$$S_k = \sum_{j=0}^{N-1} \frac{m}{\sigma_\eta^2} y(k-j) - \frac{N}{2} \frac{m^2}{\sigma_\eta^2} \tag{3-2}$$

Since S_k is Gaussian with equal variance under H_0 and H_1, we can easily compute the distinguishability metric, or signal-to-noise ratio of S_k, viz.

$$\begin{aligned} SNR\{S_{k-k_f}\} &\triangleq \sqrt{[E\{S_{k-k_f}|H_1\} - E\{S_{k-k_f}|H_0\}]^2/Var[S_{k-k_f}]} \\ &= \frac{(k-k_f)}{\sqrt{N}} \frac{m}{\sigma_\eta} \qquad \text{for } k_f \leq k \leq k_f + N \\ &= \sqrt{N} \frac{m}{\sigma_\eta} \qquad \text{for } k \geq k_f + N \end{aligned} \tag{3-3}$$

If the failure onset time were known, however, the maximum likelihood hypothesis test produces a decision statistic, R_k, from,

$$R_k = \sum_{j=k_f}^{k} \frac{m}{\sigma_\eta^2} y(j) - \frac{N}{2} m^2/\sigma_\eta^2 \tag{3-4}$$

and hence,

$$SNR\{R_{k-k_f}\} = \sqrt{k-kf}\ \frac{m}{\sigma_\eta}$$

Both $SNR\{S_{k-k_f}\}$ and $B_{01}\{R_{k-k_f}\}$ are shown in Fig. 3-2. Notice that for large values of N, (window size), one would expect considerable improvements if R_k were used instead of S_k for decisionmaking since the failure (bias) is far more distinguishable (larger B_{01}) in this case, especially for small $k-k_f$.

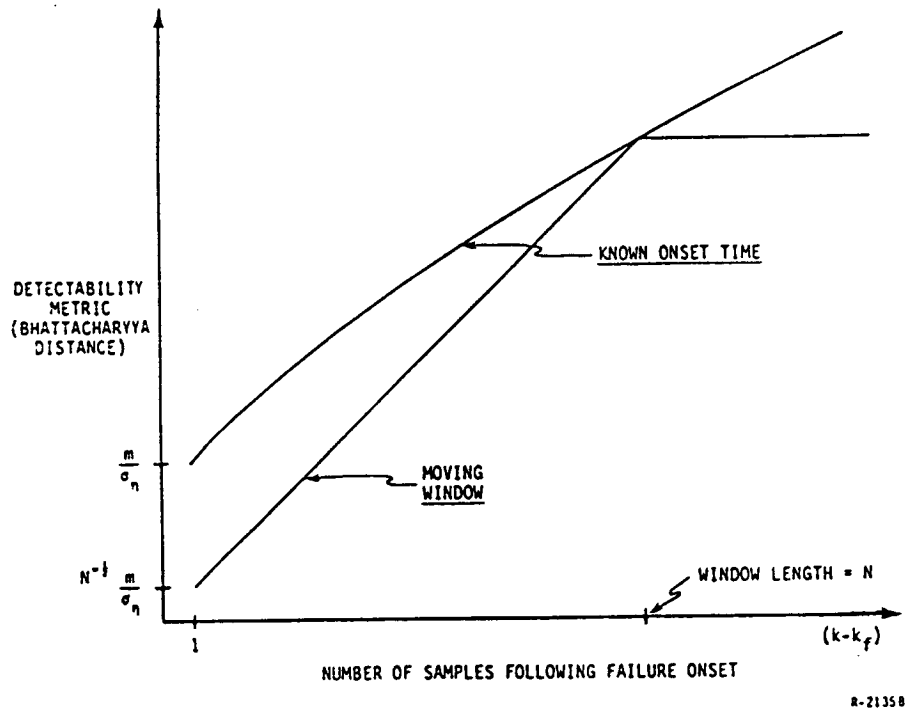


Figure 3-2. Advantage of Known Onset Time

Since, of course, the failure onset time, k_f, is unknown, alternative methods which attempt to realize performance that approaches the level obtained when the failure onset time is known have been investigated.

DECISION MECHANISMS

Several "optimal" and suboptimal methods for dealing with this issue are discussed in [17], [25]. The optimal methods, which essentially view each

sample instant as a potential failure onset hypothesis, are computationally infeasible, and many of the suboptimal algorithms which are based on these optimal methods can require extensive computational resources.

In [14] - [16] and more recently in [5], an alternative approach which avoids these complexities while sacrificing little in performance has met with considerable success. This approach is shown in Fig. 3-3. In this approach, a trigger process produces quick responding alarms on the basis of short, highly sensitive tests with full coverage of all failure modes. These alarms are then used to generate somewhat longer running, independent tests for reliably identifying the failure mode and rejecting any false start. In order to minimize the decision delay following a failure, each failure mode may have a separate trigger test. The tests which are initiated by the trigger then provide the final reliable (i.e., desired level of error probability) failure decisions. In order to reject a false start from the trigger process, several "verification" tests are triggered. These tests compare each failure mode hypothesis, (say H_j), with the no-failure hypothesis (H_0). Although any trigger may initiate the verification process, only those failure modes which have been verified will be chosen by the decision logic. If all failure modes are not verified, a false-trigger is declared. In parallel with the verify tests, several "isolation" tests are used to compare each pair of failure hypotheses which are potentially ambiguous following a given set of trigger alarms. The results of these pair-wise decisions are then combined in the decision logic to produce failure decisions.

There are several advantages to the decision structure just described. First, and most important, is that the trigger mechanism effectively provides an estimate of the failure onset time, k_f, as the beginning of the trigger

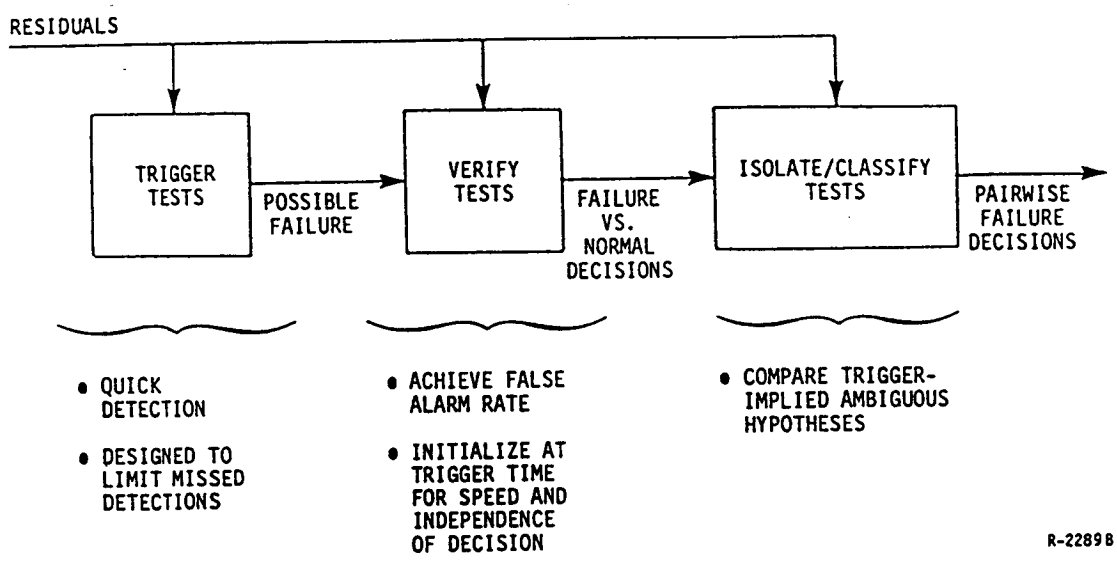


Figure 3-3. General Decision Structure

data window. This allows the use of tests based on the assumed failure onset time to be used for verification and isolation thereby achieving a high degree of decision reliability over very short time intervals. Tests such as the SPRT (a sequential decision mechanism requiring an assumed onset time) are typically very effective in realizing these advantages. In particular, these tests can easily be designed to be robust to unknown failure magnitudes in that failures which are "larger" than those considered minimal will be detected and isolated in a shorter period of time.

The second advantage of the decision structure outlined above is that typically (e.g., see [15], [5]) [6] the computational complexity of this algorithm is substantially smaller than the optimal and sub-optimal methods discussed in [17]. This is in addition to the expected performance benefits in terms of FDI robustness.

Finally, in addition to the computational advantages of the procedure described above, the partitioning of the failure space into hypothesis pairs allows us to make use of residuals in a selective way so as to minimize the decision errors and delays of each test. That is, only those residuals which provide reliable information about the hypothesis pair being tested are selected as inputs to that test.

HYPOTHESIS TEST DESIGN

The last issue we raise in regard to the design of FDI systems is the design of the various hypothesis tests which comprise the decision process. In classical hypothesis testing theory, such tests are completely defined by specification of the joint probability density function (pdf) of the sequence of residual signals. Such a characterization, however, is never completely possible since modeling errors exist, failure severity is unknown and since the inputs which excite the measurements are non-stationary and not always completely measurable. Thus, the classical theories can serve only as a starting point in the design procedure; defining a useful algorithm for hypothesis discrimination. Selection of parameters within this algorithm, however, requires a performance analysis which incorporates the uncertainties in the probabilistic description used to define the algorithm.

This process is shown graphically in Fig. 3-4. The figure emphasizes the fact that two classes of models, truth and design models, are needed; and that a variety of analysis and synthesis tools need to be developed and used in this process. For example, algorithm structures (i.e., equations for signal processing) are typically determined using a simplified design model and knowledge of various decisionmaking techniques. Of course, implicit knowledge

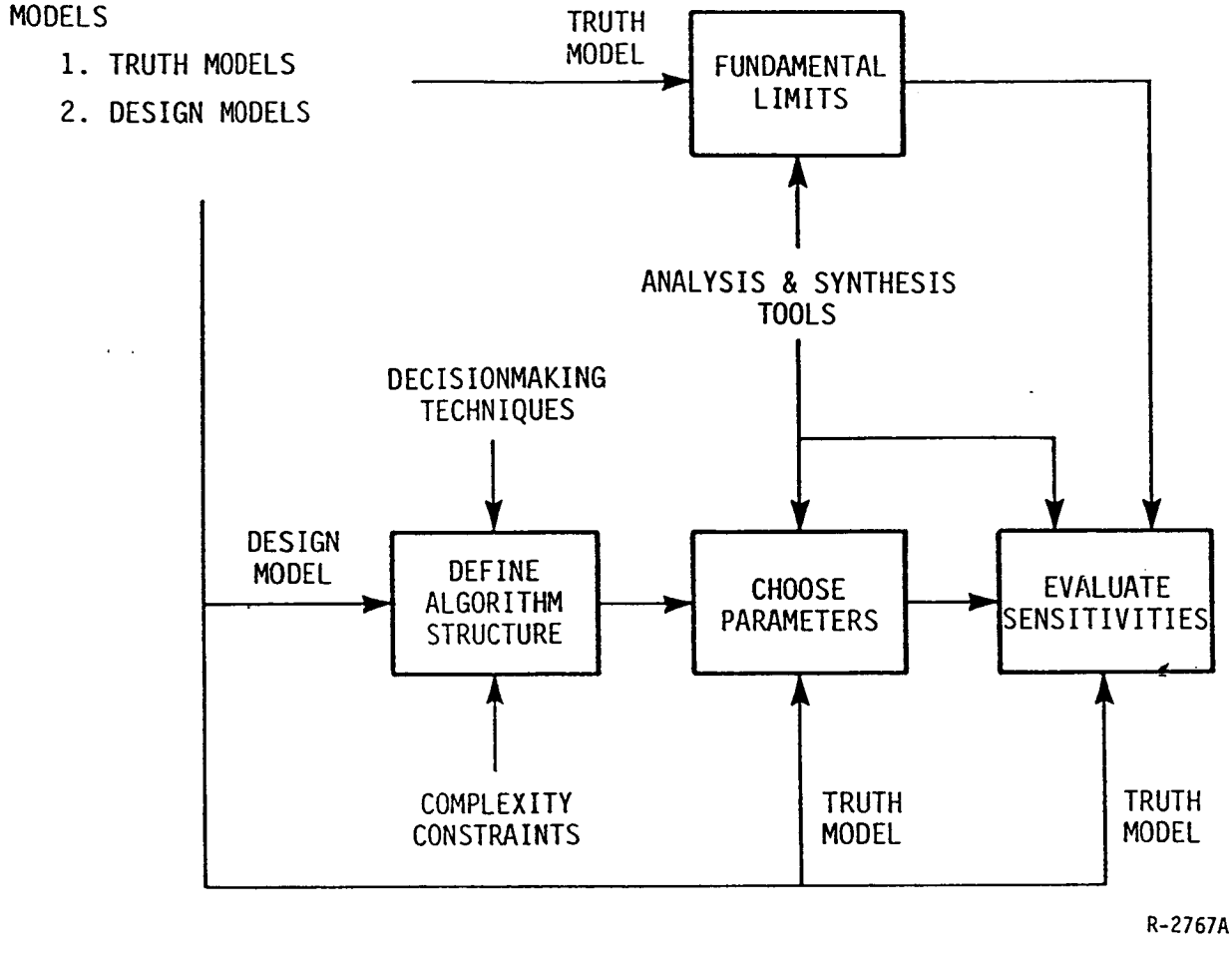


Figure 3-4. Hypothesis Test Design Process

of the truth model may also be used to ensure that the subsequent sensitivity analyses will meet the desired specifications. The truth model is then used explicitly in choosing the parameters of the decision algorithm. In order to ensure both maximal performance and robustness to off-nominal conditions, it is important that the "truth" model be a statistical model which characterizes variations in system qualities as much as possible. Finally, where such characterizations are not possible, and where optimal synthesis techniques are not

available, sensitivities to important variations must be evaluated. Iterations between such an analysis and the choice of algorithm parameters can occur and, in some cases, iteration on the algorithm structures themselves may be necessary.

Also shown in Fig. 3-4 is an evaluation of fundamental performance limits. These are really two kinds of limits of interest; one for problem feasibility and one for algorithm design. In the feasibility analysis, one asks the question How well can a decision mechanism perform under the best (but reasonable) circumstances? This topic is addressed in [8] and in Section 4 and assumes, for example, that a detailed statistical design model exists and that this model exactly describes the system behavior. If adequate performance can not be ensured for this case, the situation is most likely hopeless. For algorithm design, however, a truth model which statistically characterizes all sources of error is needed to indicate when iterations on the algorithm design are getting close to fundamental limits.

The advantage of the decentralized approach is that many of the analysis and synthesis tools required in this design process are readily available. Subsection 3.3 will detail these techniques and develop some new ones.

3.2 RESIDUAL GENERATION

As discussed in the previous section, the purpose of the residual generation function is to translate redundant information about the system (in the form of models) into signals which exhibit a well-known and easily characterized behavior (e.g., near zero with no significant trends during normal operation). These signals are formed through the relationships among measured variables embodied in the system model.

A variety of techniques for generating decentralized residuals have been developed (see [5], [22] - [24], [20]) and are described here. They are divided into 4 separate categories: memoryless relations, finite-memory relations, open-loop relations, and closed-loop relations.

3.2.1 Memoryless Relationships

Memoryless relationshps are relationships among measured variables which are valid at every time instant. They are easy to derive from static models and have been used in numerous applications from triplex or quadraplex sensor systems [6] to redundant arrays of inertial sensors [28].

In linear systems, memoryless relations are obtained as follows. The set of m measured variables, y, is related to the set of n "influence variables," x, by,

$$y = Cx \tag{3-5}$$

where C is the m x n observation matrix. Memoryless residuals or relationships are formed by solving the equation $w^TC = 0$ for all m-dimensional independent non-zero "parity check" vectors, w. The parity check vectors imply that under ideal conditions (no model error or noise), a relationship of the form $w^Ty = 0$ is valid. When a failure (e.g., sensor drift, large change in C) occurs, the residual, ν, formed using this relationship, may deviate significantly from its nominal characteristics. That is, when no failure exists

$$\nu = w^Ty = w^TCx = 0 \tag{3-6}$$

and when a failure occurs, x is something different. For example, in the case of triplex sensor systems, we have

$$C = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{3-7}$$

$$w_1 = [1, -1, 0] \tag{3-8}$$

$$w_2 = [1, 0, -1]. \tag{3-9}$$

Note that other parity check vectors satisfy $w^T C = 0$, but all are linear combinations of Eqs. 3-8, 3-9 since w_1, w_2 is a basis for the left null space of C.

Of course, the C matrix in Eq. 3-5 represents only a model of the redundancy relationships available in the static system under consideration. In order to develop residuals which are minimally sensitive to model errors and noise under normal operation, consider the uncertain system,

$$y = C_\ell x + n_\ell \tag{3-10}$$

where the observation matrix C_ℓ is parametrically related to the random variable ℓ (representing uncertainty; ℓ takes on a finite number of values in the mathematical framework of [22], n_ℓ is a zero mean Gaussian noise process with covariance R_ℓ, and x is a random variable with zero mean and covariance Σ_ℓ. From Eq. 3-10, the probability density function of y is

$$p(y) = \int p(y|\ell)\, p(\ell)\, d\ell \tag{3-11}$$

where $p(y|\ell)$ is a Gaussian density with zero mean and covariance,

$$\text{Cov}\,(y|\ell) = C_\ell \Sigma_\ell C_\ell^T + R_\ell \tag{3-12}$$

In reference [24], several methods of generating robust residuals using this formulation are given. They include minimization of the variance of the residuals, minimizing the average entropy and several methods which make explicit use of failure models to guarantee sensitivity to particular failures as well as insensitivity to model error and noise. In minimizing the variance of the residuals, for example, we define a t-dimensional residual vector, ν;

$$\nu = W^T y \tag{3-13}$$

The function to be minimized is $J=\text{Trace}[E\{\nu^T\nu\}]$ which can be written

$$J = \text{Tr}\{W^T[\int (C_\ell \Sigma_\ell C^T_\ell + R_\ell)\ p(\ell)\ d\ell]W\}$$

$$= \text{Tr}\{W^T\ \overline{C}\ W\} \tag{3-14}$$

where $\overline{C}$ is used to denote the term in brackets. If we constrain W so that it is non-zero, (e.g., $W^TW = I$), then it can be shown that since $\overline{C}$ is symmetric, the solution is to take the columns of W as the eigenvectors corresponding to the t-smallest eigenvalues of C.

One can interpret this result geometrically by considering the following example. Let $\Sigma_\ell=I$ and $R_\ell=0$ for all ℓ and let

$$C_\ell = \begin{bmatrix} \cos(\ell) \\ \sin(\ell) \end{bmatrix} \tag{3-15}$$

where ℓ is a uniform random variable on $[\theta_1, \theta_2]$. In this case $\overline{C}$ can be computed and the eigenvector corresponding to the smallest eigenvalue computed. Through a change of variables, it can be shown that the "optimal" residuals are computed by projecting y onto a space which is orthogonal to the vector,

$$\theta = \begin{bmatrix} \cos \dfrac{(\theta_1 + \theta_2)}{2} \\ \sin \dfrac{(\theta_1 + \theta_2)}{2} \end{bmatrix} \tag{3-16}$$

This situation is depicted in Fig. 3-5:

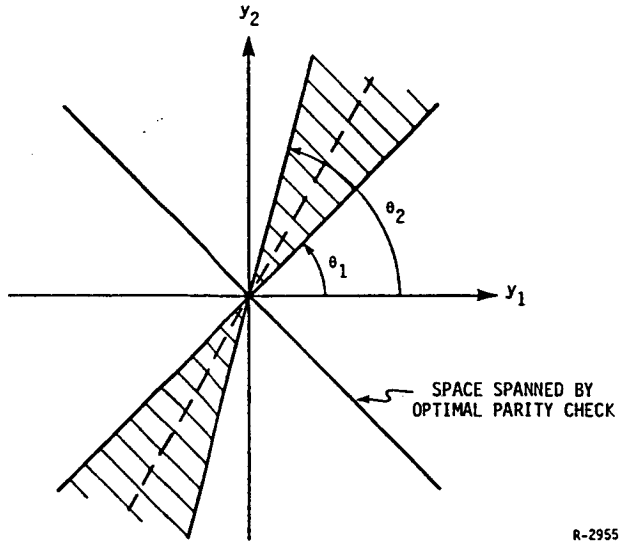


Figure 3-5. Geometrical Interpretation of Robust Residuals

If the linear relationship of Eq. 3-5 is not available, it may still be possible to generate memoryless residuals. Suppose

$$y_1 = f(y_2) \tag{3-17}$$

models a static relationship between measured variables y_1 and y_2. Then the residual,

$$\nu = y_1 - f(y_2) \tag{3-18}$$

can be computed at each time instant and used for FDI.

As an example of this kind of memoryless relation, consider the force balance relationship in a rigid aircraft model. The aerodynamic forces on an aircraft can be related to the relative velocity of the aircraft with respect to the air mass, V, the angular velocity of the aircraft about its C.G. (ω) and the deflections of the control surfaces, δ, by

$$F = f(V, \omega, \delta) \tag{3-19}$$

This force, in turn, can be directly related to the output of the compensated (for off C.G. effects) accelerometer measurements. That is, since the aircraft obeys [29],

$$m(\dot{V} + \omega \times V) = \underline{F} + m\underline{g} \tag{3-20}$$

and the accelerometer readings (which measure specific force) obey,

$$A_m = \dot{V} + \omega \times V - \underline{g} \tag{3-21}$$

a set of three translational residuals can be defined by,

$$\nu = m A_m - f(V, \omega, \delta) \tag{3-22}$$

when air data, inertial data, and surface deflections are all measured. When the aircraft is operating normally, sensor errors and errors in the aero-model cause these residuals to deviate from zero. If a control derivative changes or a measured deflection differs from an actual (or effective) deflection, these residuals can deviate significantly from their behavior under normal conditions, and are, therefore, useful for detecting such failure modes.

3.2.2 Finite Memory Relationships

A finite memory relation is one in which measured variables over a finite window in time are used. In the linear case, they can be derived similarly to memoryless relations as follows. Consider the linear system

$$x(k+1) = Ax(k) + Bu(k)$$
$$y(k) = Cx(k) + Du(k) \tag{3-23}$$

where

$x(k)$ = NS - dimensional state vector.

$y(k)$ = NO - dimensional output vector.

$u(k)$ = NC - dimensional input vector.

A redundancy relationship is now defined as a linear combination of measurements and controls over a finite window of observation. Specifically, if we let $Y_p^T(k) = \left(y^T(k),\ y^T(k+1),\ \ldots\ y^T(k+p)\right)$, and $U_p^T(k) = \left(u^T(k),\ u^T(k+1),\ \ldots\ u^T(k+p)\right)$, then redundancy relationships take the form

$$\nu(k) = W^T \begin{bmatrix} Y_p(k) \\ U_p(k) \end{bmatrix} = W_y^T\, Yp(k) + W_u^T U_p(k) \tag{3-24}$$

where $\nu(k)$ is the t-dimensional residual vector which, under ideal circumstances (no noise or modeling error) is identically zero, and W is the parity check matrix. Next, we can expand $Y_p(k)$ in terms of the system matrices (A,B,C,D,) as (see [24])

$$Y_p(k) = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^p \end{bmatrix} x(k) + \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & & \\ \vdots & & & \\ CA^{p-1}B & \cdots & & CB\ D \end{bmatrix} U_p(k) \tag{3-25}$$

or

$$Y_p(k) = M_p\, x(k) + N_p\, U_p(k) \tag{3-26}$$

Thus the residual ν, Eq. 3-24, can now be written as,

$$\nu(k) = W^T \begin{bmatrix} M_p & N_p \\ 0 & I \end{bmatrix} \begin{bmatrix} x(k) \\ U_p(k) \end{bmatrix} \tag{3-27}$$

When no modeling error or noise is present, we can make $\nu(k)$ identically zero by choosing W as an orthogonal basis for the left null-space of the matrix,

$$\overline{M}_p \triangleq \begin{bmatrix} M_p & N_p \\ 0 & I \end{bmatrix} \tag{3-28}$$

That is, we find all the vectors for which $w^T \overline{M}_p = 0$ and form the parity check matrix using these vectors for its rows.

Comments

1. The minimum number of independent parity checks for any p is NO(p+1)-NS when NO(p+1) > NS.

2. As discussed in [30], one need only look at values of p=0, ... , NS to find all of the independent parity checks.

3. The solution for W can also be obtained by finding the vectors which satisfy $W_y^T M_p = 0$, and then solving $W_y^T N_p + W_u^T = 0$ where $W^T = (W_y^T, W_u T)$.

Uncertainty and noise can be added to the system model of Eq. 3-23 as in the static (memoryless) case and similar results derived [24]. The window length, p, however, in this case is not as easily determined.

In addition to the linear case above, finite memory relationships can also be obtained for general (nonlinear) system dynamics when sufficient state

measurements are available. Consider the nonlinear time-invariant, discrete-time system model,

$$x_{k+1} = f(x_k, u_k) \tag{3-29}$$

$$y_k = x_k + v_k \tag{3-30}$$

$$z_k = u_k \tag{3-31}$$

Finite-memory nonlinear relationships between the measured variables (y_k and z_k) are then, obviously, given by

$$y_{k+1} = f(y_k, z_k) \tag{3-32}$$

The residual, $\nu = y_{k+1} - f(y_k, z_k)$, can then be generated.

Several comments about these relationships and their associated residuals are pertinent at this point. First, note that although Eq. 3-32 is written in vector form, each component may be considered as a separate relationship and evaluated in terms of its usefulness for FDI. Furthermore, full state measurement is not necessary for generating those individual relationships in Eq. 3-32 which only depend on a subset of states.

As an example of nonlinear finite memory relations, consider the so-called "rotational" residuals used in [5] and [14]. In order to form rotational residuals, we first write a discrete time, nonlinear, time invariant state space model for the aircraft; e.g.,

$$x(k+1) = a\big(x(k)\big) + b\big(u(k)\big) \tag{3-33}$$

where x(k) is the n-dimensional state vector and u(k) is the m-dimensional input vector. Moment balance and rotational dynamic relationships give rise to three components of Eq. 3-33 which correspond to the angular velocity

states, $\underline{\omega}(k)$. Since we have sufficient measurements of the states, x, and inputs, u, for these three equations, three rotational residuals can be defined by,

$$\nu = \begin{pmatrix} \nu_p \\ \nu_q \\ \nu_r \end{pmatrix} = \omega(k+1) - a_\omega(x_m(k)) - b_\omega(u_m(k)) \qquad (3\text{-}34)$$

where $a_\omega(\cdot)$ and $b_\omega(\cdot)$ are the components of $a(\cdot)$ and $b(\cdot)$ which correspond to the states, ω, and where x_m and u_m are the required state and control measurements.

3.2.3 Open Loop Relationships

If we have insufficient state measurements for individual finite memory relationships, we may derive residuals in an open loop manner as follows. Consider the general non-linear system model of Eq. 3-29. An open-loop estimate of x_k may be obtained from the recursive procedure,

$$\hat{x}_{k+1} = f(\hat{x}_k, z_k) \qquad (3\text{-}35)$$

If the measurements are given by

$$y_k = h(x_k) + v_k \qquad (3\text{-}36)$$

$$z_k = u_k \qquad (3\text{-}37)$$

then a residual vector may be formed from

$$\nu(k) = y_k - h(\hat{x}_k) \qquad (3\text{-}38)$$

Comments:

1. The residuals in 3-38 are only considered as decentralized if f and h represent a decoupled subset of a complete system model, (e.g., an actuator model).

2. In order for Eqs. 3-35 - 3-38 to be useful, the system Eq. 3-35 must be stable so that initial condition errors do not cause serious deviations of the residual from zero.

As a simple example, consider a first order linear system representing, possibly, an actuator model. Let,

$$\hat{\delta}_A(k) = \hat{x}(k)$$

$$\hat{x}(k+1) = \alpha\, \hat{x}(k) + (1-\alpha)\, \delta_c(k) \qquad (3\text{-}39)$$

where

$\hat{\delta}_A$ = estimated surface deflection

$\hat{x}$ = estimator state variable

δ_c = commanded deflection

The residual, ν, is the difference between the measured surface deflection, δ_A, and the estimated deflection, $\hat{\delta}_A$. That is,

$$\nu(k) = \delta_A(k) - \hat{\delta}_A(k) \qquad (3\text{-}40)$$

The linear model in Eq. 3-39 is typically chosen to match the DC gain and low frequency phase of the true actuator. In addition, rate and position limits can easily be added to the residual generation procedure by appropriate modifications of the estimate, $\hat{\delta}_A$, at each stage.

3.2.4 Closed Loop Relationships

This case is the most general and is equivalent to the Kalman filter (or extended Kalman Filter) approach if no use is made of the natural system decoupling. In general, if the system and observation model are given by Eqs. 3-29 and 3-36, 3-37 respectively, then a closed-loop residual can be formed from,

$$\bar{x}_{k+1} = f(\bar{x}_k + G\nu_k,\ z_k) \tag{3-41}$$

$$\nu_k = y_k - h(\bar{x}_k) \tag{3-42}$$

Clearly, the open loop case Eqs. 3-35 - 3-40 is equivalent to the closed loop case with G = 0. Also, when the functions $f(\cdot,\cdot)$ and $h(\cdot)$ are linear, it is possible to relate the set of finite memory relationships to the closed loop relationships as well. This can be seen in the one dimensional linear full state measurement case as follows.

Let the system be described by,

$$X_{k+1} = AX_k + B\mu_k \tag{3-43}$$

The one dimensional closed loop residual, ν_{CL}, is given by

$$\bar{X}_k = A[\ \bar{X}_{k-1} + G(X_{k-1} - \bar{X}_{k-1})\ ] + B\mu_{k-1} \tag{3-44}$$

$$\nu_{CL}(k) = X_k - \bar{X}_k \tag{3-45}$$

Obviously, when G=0 we get the open loop case, and when G=1 we get the finite memory case. Furthermore, rewriting Eq. 3-45 in terms of the state and estimate at time k-1, we can derive the following relationships;

$$\nu_{CL}(k) = -AG\ \nu_{CL}(k-1) + \nu_{OL}(k) \tag{3-46}$$

$$\nu_{CL}(k) = \nu_{FM}(k) + A(1-G)\ \nu_{CL}(k-1) \tag{3-47}$$

where $\nu_{FM}(k)$ denotes the finite memory residual (G-1) and $\nu_{OL}(k)$ denotes the open loop residual (G=0).

In general, closed loop residuals ($G \neq 0,1$ in the above) will have failure effects which appear in all the residuals in a fairly complex manner. Furthermore, the affects of modeling error on closed loop residuals are not easily characterized [6]. Thus, decentralization of the closed loop residual generation process is possible only when the system has naturally decoupled modes.

Finally, we note that the choice of Kalman gain, G, for FDI purposes, is not necessarily straightforward. The ideal choice of G would produce residuals which are insensitive to model error and noise, and respond quickly to failures, with high sensitivity, in a well-defined and robustly distinguishable manner, for each class of failure types. Although this is a tall order, if we take advantage of system decoupling, we may reduce some of the requirements on G by ensuring that only a subset of failures appear in the residuals generated by Eqs. 3-41, 3-42 and only a subset of model errors affect the behavior of ν_k under normal circumstances. Some recent work in the area of robust Kalman filter design for FDI is given in [31].

3.2.5 Summary of Residual Generation Issues

We have seen that decentralization of the residual generation process consists in defining residuals in such a way that both failures and model errors affect only small subsets of residuals. This will, in turn, allow us to make selective use of the redundancy relations in the decision process. Only those relations or residuals which provide significant information about a specified set of hypotheses need to be considered when defining algorithms for distinguishing those hypotheses.

All of the residuals we have defined are, of course, based on design models. The complexity of the residual generation process is, therefore,

directly related to the complexity of these models. In some cases, it may be appropriate to use a combination of residual generation techniques to form a suitable set of individual or decentralized residuals. Whatever the process, however, it is important to recall that we will always use a model of reality and, in so doing, we require that an analysis of modeling error be made before any residual is used in the decision process.

3.3 DECISION PROCESS DESIGN

In subsection 3.1, we described a general structure for the information collection part of the decision process. This structure consists of a two-level process with three functional blocks: trigger, verify, and isolate (see Fig. 3-3). Each block is composed of a variety of statistical tests which are designed independently to make reliable decisions and make use of only the best relationships in each particular test.

Figure 3-4 illustrated the basic methodology behind the design of these tests. In this subsection we will develop, largely by example, some of the most useful analysis and synthesis tools which comprise this methodology. First, however, a short review of concepts in statistical hypothesis testing is given.

3.3.1 Review of Statistical Hypothesis Testing

A great deal of literature has appeared throughout the years on the subject of statistical hypothesis testing. Nevertheless, many results are based on the same fundamental concepts involving probability theory and Gaussian statistical assumptions. These concepts are now described.

UNITARY DECISIONMAKING

The first concept typically appears in the statistics literature (e.g., see [B-stat]) and involves the problem of deciding whether or not an observed set of samples (measurements) can be described by some assumed underlying probability distribution. What is generally true, however, is that given this assumed distribution, one can only reliably reject this "hypothesis." Hence, in the statistics literature, one finds referral to forming a "null hypothesis" in which all hypotheses besides the one we really wish to accept are modeled by a single distribution. If the null hypothesis is reliably rejected, then its complement can be reliably accepted.

To see why hypotheses can only be reliably rejected, consider the case where a random vector, y, is assumed to have a jointly normal distribution under the null hypothesis, H_o. Denote this by $y \sim N[m; \Sigma]$ corresponding to the probability density function (pdf)

$$p_y(Y) = \frac{1}{[2\pi|\Sigma|]^{n/2}} \exp -\frac{1}{2}(Y-m)^T \Sigma^{-1}(Y-m) \qquad (3\text{-}48)$$

where n is the dimension of the random vector, y, and $|\cdot|$ denotes the determinant function.

We now wish to define a decision region, D, in which the condition $y \in D$ is very unlikely. Figure 3-6 shows one definition of D for the zero/mean one dimensional case (note that D is not unique). The area under the shaded portion of the curve is α. For multivariable zero-mean unimodel densities, we can choose D as follows:

$$D = \{Y: \Pr[\|y\|^2 > \|Y\|^2] \leq \alpha\} \qquad (3\text{-}49)$$

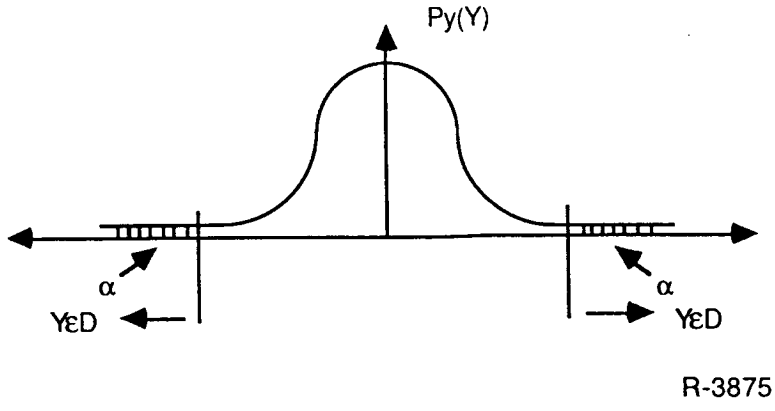


Figure 3-6. One Dimensional Rejection Region

where $\|\cdot\|$ denotes Euclidean norm (others are possible) and where α specifies the level of significance. When the random variable y is used to represent a residual vector and the norms in Eq. 3-49 are taken with respect to some covariance matrix, Σ, then the rejection criterion $Y \varepsilon D$ is equivalent to the well known weighted-sum-of-squared-residuals (WSSR) test.

In the FDI process, rejection tests are useful in the trigger process (which monitors operation). However, since no mention of alternate hypotheses (failures) is made, there is no guarantee that the rejection test of Eq. 3-49, if used as a trigger, would be sensitive to all important failures. In fact, it is certain that this test is <u>not</u> maximally sensitive to all failure modes. Thus, in failure detection, since alternate hypotheses are sometimes available, we need to consider binary, or in general, M-Ary decisionmaking.

M-ARY DECISIONMAKING

The general problem here can be described as follows. Consider an observed signal y(k) which can be characterized in a probabilistic sense by two probability density functions (pdf's) each being valid under two different hypotheses. (Only 2 pdf's are considered here for simplicity, however, multiple hypotheses may also be considered). That is, we can define two conditional pdfs, $p(Y|H_0)$ and $p(Y|H_1)$, where Y is any specified set of the signal y(k), (e.g., $y=\{y(k);\ k=k_0, \ldots, k_1\}$).

The problem which must be solved is the definition of decision regions D_i which map the observables, Y, into decisions about the system status. That is, we will decide that H_i is true when $Y \in D_i$, i=0, 1.

In order to choose D_i in an "optimal" manner, we must specify an optimization problem. Two commonly mentioned problems are the Bayesian and Neyman-Pearson problems both of which are now described.

Bayesian Hypothesis Testing - In this method we attempt to choose D_i through minimization of an expected cost function associated with the four different decision regions shown in Fig. 3-7. The total cost C is given by

$$C = \sum_{i,j} C_{ij} \cdot Pr\ (Y \in D_i | H_j) \qquad (3\text{-}50)$$

Where C_{ij} is the cost of deciding that hypothesis i is true when in fact hypothesis j is true (i.e., $Y \in R_{ij}$). The optimization problem,

$$\min_{D_i, i=0,1} E\{C\}$$

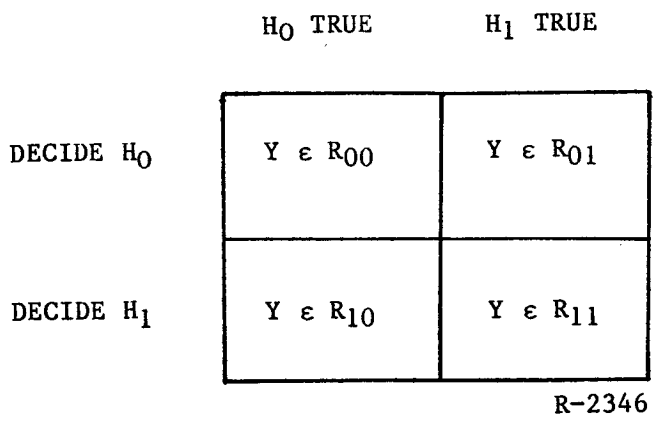


Figure 3-7. Decision Regions

then, has the solution [32],

$$D_1 = \quad Y: \quad \frac{P(Y|H_1)}{P(Y|H_0)} \quad > \quad \frac{P_0\,(C_{10} - C_{00})}{P_1\,(C_{01} - C_{11})}$$

$$D_0 = \quad Y: \quad Y \notin D_1(Y) \qquad (3\text{-}51)$$

where P_1 and P_0 are the a-priori probabilities of each hypothesis. Notice that the form of the solution involves the comparison of the ratio of two pdf's or "likelihood ratio", to a threshold which is a function of the costs of various decision conditions. The form of this solution is quite general as it is also the form of the solution to the Neyman-Pearson problem described below.

Neyman-Pearson Formulation - This method recognizes that the Bayesian costs are difficult to specify and formulates a problem in terms of specific performance traits. In particular, let P_{FA} denote the probability of incorrectly deciding that H_1 is true when H_0 is, in fact, true. Also, let P_D denote the probability that the correct decision, H_1, is made when H_1 is true. The optimization problem,

$$\max P_D$$

$$\text{subject to } P_{FA} = \gamma \tag{3-52}$$

then, has the solution,

$$D_1 = \left\{ Y: \frac{P(Y|H_1)}{P(Y|H_0)} > t \right\}$$

$$D_0 = \{Y: Y \notin D_1(Y)\} \tag{3-53}$$

where we have,

$$P_{FA} = \int_{D_1} P(Y|H_0)\, dY \tag{3-54}$$

$$P_D = \int_{D_1} P(Y|H_1)\, dY \tag{3-55}$$

Thus the form of the solution to this problem is the same as the Bayesian case. The threshold, t, must be determined so that Eq. 3-54 is satisfied, and the resulting performance, P_D, is determined from Eq. 3-55. The tradeoff between P_{FA} and P_D is typically expressed by a graph called the receiver operating characteristic (ROC) as shown in Fig. 3-8. At the extremes we see that:

1. if $t = \infty$, then $P_{FA} = 0$ and $P_D = 0$, and
2. if $t = 0$, then $P_{FA} = 1$ and $P_D = 1$.

Of course, neither extreme is useful and the choice of operating point depends on some idea of acceptable performance. When this idea is expressed in terms of the Bayesian cost, the operating point is determined by the threshold choice of Eq. 3-51.

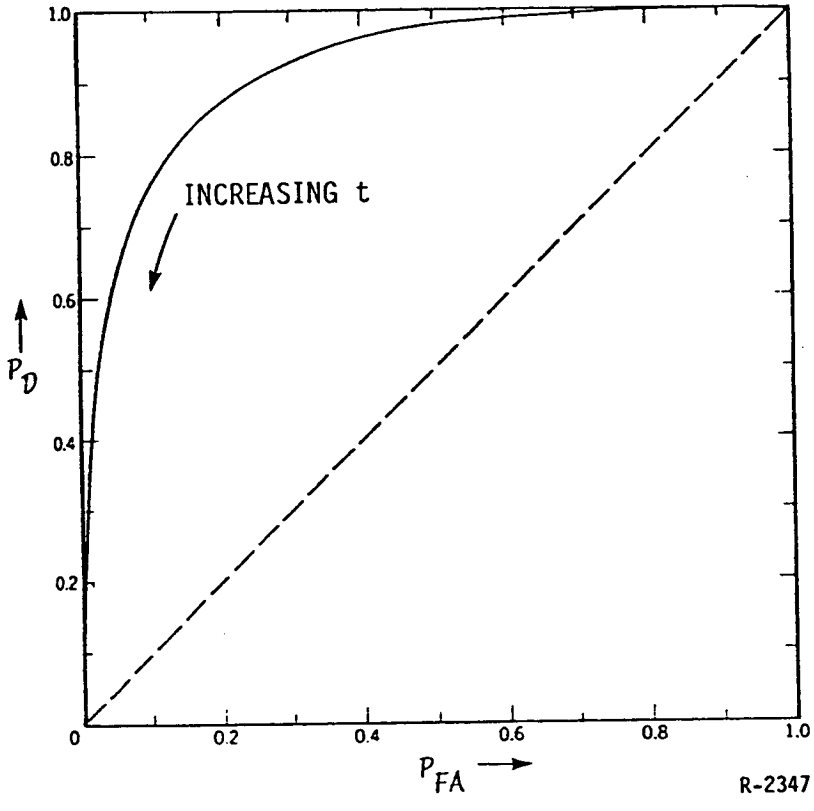


Figure 3-8. Receiver Operating Characteristic (ROC)

Example - As an example of the above concepts, consider a situation in which under H_0, y_k is a white gaussian noise process while under H_1, y(k) is the same noise process plus a constant. That is,

$$H_0: \quad y_k = \eta_k$$
$$H_1: \quad y_k = \eta_k + m \tag{3-56}$$

Now, if decisions are made on the basis of a fixed set of samples of y(k), say $y = \{y(k): \ k=1, 2, \ldots, N\}$, the decision rule becomes

$$\ell = \sum_{j=1}^{N} m^T \Sigma^{-1} (y_j - m/2) \underset{D_0}{\overset{D_1}{\gtrless}} t^1 \tag{3-57}$$

where Σ is the covariance matrix of the white noise process and $t^1 = \ell n(t)$ since Eq. 3-57 is the natural logarithm of the decision rule of Eq. 3-53. The decision statistic, ℓ, is called the log-likelihood ratio (LLR) and Eq. 3-57 is an example of an "LLR test". We also note at this point that if Eq. 3-57 operates over a moving window of data, y(k), as is frequently done in practice, then the LLR decision statistic can, more generally, be viewed as the output of a finite impulse response filter. This "filter" interpretation is useful from the standpoint of complexity reduction since similar filters with fewer states can be used and analyzed.

Returning to our example above, we wish to characterize the performance of the test in Eq. 3-57. Since Y is jointly Gaussian, the decision statistic ℓ is a Gaussian random variable, and can, therefore, be characterized in terms of its mean, $\bar{\ell}$, and variance, σ_ℓ^2 under each hypothesis. These are given by

$$H_0: \quad \bar{\ell}_0 = -Nm^T\Sigma^{-1}m/2 < 0$$
$$\sigma_\ell^2 = Nm^T\Sigma^{-1}m \qquad (3\text{-}58)$$

$$H_1: \quad \bar{\ell}_1 = -\bar{\ell}_0 > 0$$
$$\sigma_\ell^2 = Nm^T\Sigma^{-1}m \qquad (3\text{-}59)$$

The probability of false alarm P_{FA} can be computed as a function of t^1 using Eq. 3-58 and percentage points of the Gaussian distribution. If we let $Q(\zeta)$ represent the area under the zero mean unit normal function from (ζ, ∞), then we have

$$P_{FA} = Q\left(\frac{t^1 - \bar{\ell}_0}{\sigma_\ell}\right) \qquad (3\text{-}60)$$

Furthermore, P_D may be computed from

$$P_D = Q\left(\frac{t^1 - \bar{\ell}_1}{\sigma_\ell}\right) \tag{3-61}$$

Eqs. 3-60 and 3-61 completely specify the ROC. If we wish to make $P_{FA} = 1-P_D$ (i.e., equal costs for both types of errors), it can be shown [32] that $t^1 = 1/2\ (\bar{\ell}_1 + \bar{\ell}_0) = 0$ resulting in performance which is completely characterized by the quantity $d^2 = (\bar{\ell}_1 - \bar{\ell}_0)^2 / \sigma_\ell^2$ (Note that d^2 is precisely the signal-to-noise ratio used in Eq. 3-3). That is, larger values of d imply smaller probabilities of incorrect decisions. Furthermore, it can be shown [32] that, for any t^1, P_D can be determined from d and P_{FA}, with larger values of d corresponding to larger values of P_D. This is shown in Fig. 3-9. For the example we are considering here,

$$d^2 = N\ m^T \Sigma^{-1} m \tag{3-62}$$

Thus, we see that tradeoffs exist between performance (as expressed by d^2) and window length, N, failure size, m, and noise covariance, Σ.

The Sequential Probability Ratio Test - The LLR tests described above are known as fixed sample size tests for obvious reasons. Another test which has been used extensively for FDI is the sequential probability ratio test (SPRT) [5] [33]. Rather than basing decisions on a fixed sample of data, the SPRT decides automatically when enough samples have been taken to make a reliable decision. That is, given P_{FA} and P_D, the SPRT chooses one of three decisions after each sample: H_0 is true, H_1 is true, or take another sample. These choices are completely defined in terms of an LLR statistic and two thresholds by

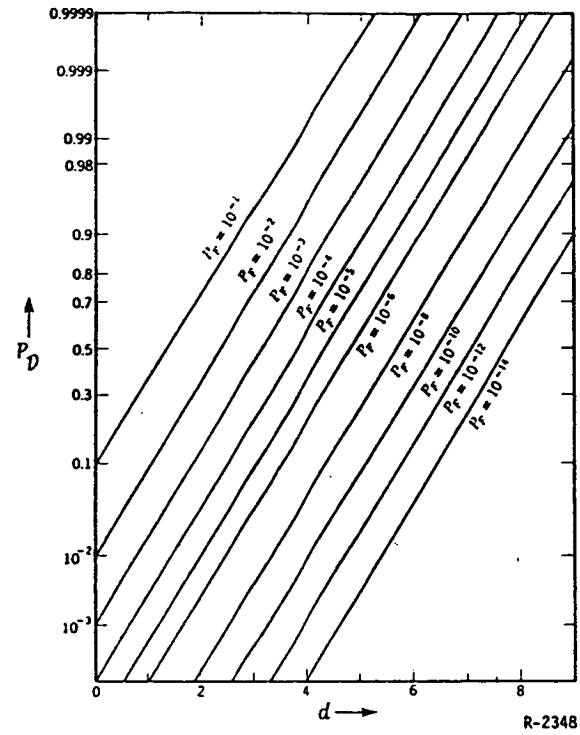


Figure 3-9. Detection Probability Versus d

$$\begin{aligned} \ell_k > t^+ &\Longrightarrow \text{decide } H_1 \\ \ell k < t^- &\Longrightarrow \text{decide } H_0 \\ t^- < \ell_k < t^+ &\Longrightarrow \text{take another sample} \end{aligned} \tag{3-63}$$

where $\ell_k = \ell n\,[p(Y_k|H_1)/p(Y_k|H_0)]$ and $Y_k = \{y(j):\ j = 1,2\ \ldots,\ k\}$. Under ideal circumstances, the test is guaranteed to terminate and is the optimal test in the sense of minimizing the number of samples [33]. Also, if we choose t^+ and t^- by

$$\begin{aligned} t^+ &= \ell n \left[\frac{1-\beta}{\alpha}\right] \\ t^- &= \ell n \quad [\beta/1-\alpha] \end{aligned} \tag{3-64}$$

then it can be shown that [34] $P_{FA} < \alpha$ and $P_D > 1-\beta$.

DISCUSSION

This concludes the review of basic hypothesis testing theory. The results presented show that in many cases of practical interest, the hypothesis tests which comprise the decision mechanism part of the FDI process consist of the formation of a decision statistic (e.g., the log likelihood ratio) and the comparison of that statistic with one or more thresholds. Optimal computations of the parameters of the mechanism for generating decision statistics and the thresholds were given on the assumption that the underlying pdfs accurately described the behavior of the observed quantities.

As we have argued in subsection 3.1, however, it is rarely possible to adequately characterize these pdfs in practice. Therefore, the design of hypothesis tests can use these classical ideas only as a means for determining algorithm structure. The parameters, however, must be chosen not for optimal performance in the nominal case, but for robust performance, i.e., maximal performance when averaged over all sources of modeling error. In the next subsection we go into greater detail on these ideas.

3.3.2 Decision Process Design and Analysis

As discussed in subsection 3.1, the classical techniques discussed above form only a basis for defining the structure of the hypothesis tests which comprise the trigger, verify, and isolate phases of the decision process. The parameters of these algorithms, however, must be selected so that the resulting tests are robust to errors in the statistical characterizations which were used to develop them. The tools which will be developed in this section are based on the notion that these tests can be "robustified" by selecting their parameters to "optimize" and/or tradeoff desired performance measures with respect to a statistical truth model. That is, rather than using the

parameters specified applying classical theories to the design model, we wish to select parameters which optimize performance and achieve the desired tradeoffs when averages are taken with respect to a truth model that includes all error sources.*

To be more explicit, we assume in this methodology that the truth model can be described by two pdfs for every hypothesis. The first pdf is the conditional distribution of the measured quantities conditioned on a random vector which represents the sources of error which are ignored in the design model. This is denoted by $p(y\theta, H_i)$. The second pdf is a characterization of the model error vector, θ, $p(\theta|H_i)$. The design model is based on the assumption that θ takes on a specific value (usually zero). Thus, the structure of the decision algorithms are determined using $p(y|\theta=\theta_o, H_i)$. The algorithm so determined will have a number of parameters, generically denoted by P. To select P, we go back to the original hypothesis testing algorithm (e.g., max P_D subject to constant P_{FA}) and select P using the truth pdfs,

$$p(y|H_i) = \int p(y|\theta, H_i)\, p(\theta|H_i)\, d\theta \qquad (3\text{-}65)$$

Although this concept handles many sources of uncertainty, there may be other unknown quantities which can not be adequately represented by a random variable with some known distribution (e.g., failure magnitude). In this case we must look at the sensitivities of the various performance measures to variations in the unknown quantities.

*A natural question arises here: Why not design everything for the truth model? The answer is that one would then expect good performance only for the truth model and little robustness to its assumptions. The truth model is not reality either. It is just a vehicle for obtaining a warm feeling about the robustness of the algorithm based on the design model.

EXAMPLE 1 (DETECTION)

This example deals with fixed sample size tests. It illustrates the use of the classical theories to define an algorithm structure and the use of a truth model to select parameters of that structure based on optimization and engineering tradeoffs. This example, while very simple, will be very useful in designing trigger tests for aircraft path failures.

The design model for this example is given by Eq. 3-56. Under H_o the observed vector process is white Gaussian noise, and under H_1 this noise is contaminated by a constant vector. The decision process structure for a fixed sample size, N, is given in Eq. 3-57. Since we wish to select the parameters of this algorithm using a truth model, we rewrite 3-57 as

$$S = P^t \sum_{j=1}^{N} y_j \underset{D_o}{\overset{D_1}{\gtrless}} t'' \tag{3-66}$$

Thus, the parameters which must be chosen (given N) are a projection vector, P, and the threshold, t". The truth model we will use to select these parameters is given by

$$H_o : y_k = b + n_k \tag{3-67}$$

$$H_1 : y_k = b + n_k + m \tag{3-68}$$

where b is a random zero mean Gaussian constant vector whose covariance matrix is Σ_b, n_k is a zero mean white noise process with covariance matrix, Σ_n, and m is a known constant vector. As discussed in subsection 3.3.1, the performance of the test in 3-66 can be completely characterized by the distinguishability

metric d^2, however, we now compute d^2 using the truth model of Eqs. 3-67 and 3-68. In particular we can compute the mean and variance of S under H_o and H_1 (denoted by $\bar{S}_o$, $\bar{S}_1$, σ_o^2, and σ_1^2) as

$$\bar{S}_o = 0 \tag{3-69}$$

$$\sigma^2 = \sigma_1^2 = \sigma_0^2 = P^t(N^2\,\Sigma_b + N\,\Sigma_\eta)P \tag{3-70}$$

$$\bar{S}_1 = N\,P^t m \tag{3-71}$$

This gives,

$$d^2 = \frac{N(P^t m)^2}{P^t(N\Sigma_b + \Sigma_\eta)P} \tag{3-72}$$

Recalling that large values of d^2 result in large probabilities of detection and low probabilities of false alarm, we wish to choose P to maximize the right hand side of Eq. 3-72. To do this, define the following quantities,

$$\bar{C} = N\,\Sigma_b + \Sigma_\eta, \tag{3-73}$$

$$\bar{Q}^t\,\bar{Q} = \bar{C}, \tag{3-74}$$

$$\bar{P} = \bar{Q}\,P, \tag{3-75}$$

$$\bar{m} = \bar{Q}^{-t}\,m. \tag{3-76}$$

This allows us to rewrite Eq. 3-72 as

$$d^2 = N(\bar{P}^t\,\bar{m})^2 \,/\, (\bar{P}^t\,\bar{P}) \tag{3-77}$$

which is clearly just the square of the magnitude of the projection of $\bar{m}$ onto $\bar{P}$. Thus d^2 is maximized when $\bar{P} = K\bar{m}$ (K is any real scalar) and using Eqs. 3-73 through 3-76 we can solve for P as

$$P = K \overline{C}^{-1} m \tag{3-78}$$

Next the threshold t" is chosen. We recall that the structure of Eq. 3-66 could be derived from the problem of maximizing P_D subject to a given value for P_{FA}. To achieve the desired P_{FA} we must compute it as a function of t" using the truth model. Since this example deals with Gaussian statistics, this computation is straightforward. In particular,

$$P_{FA} = Q\ (t"/\sigma) \tag{3-79}$$

where σ is defined in Eq. 3-70 and $Q(\cdot)$ is the error function discussed in subsection 3.3.1. The threshold t" is then chosen using percentage points of the Gaussian density. For example, to achieve $P_{FA} = 10^{-4}$ we need t" approximately equal to 3 σ.

Similarly, for a known value of m, we can compute P_D using Eq. 3-61 and the truth model pdfs. When 'm' represents a failure, however, it is sometimes of interest to define a minimal failure in terms of a <u>desired</u> value of P_D. This can also be done using Eq. 3-61 and the truth pdfs. For example, to achieve $P_D = 10^{-4}$, we need $(\overline{S}_1 - t") = 3\ \sigma$. Furthermore, since d^2 (and, therefore, P_D) is monotonic in $\overline{S}_1$, we can guarantee that any failure magnitude which is larger than the minimal one defined in this way will achieve a P_D which is no smaller than the desired value. This is a very desirable trait for failure detection systems in which the size of a failure is infrequently known a priori.

Another tradeoff which can be accomplished using the above computations is the choice of sample size, N. Since all the computations are a function of N, we could, for example, proceed as follows:

1. Vary N
2. Choose P to optimize d^2
3. Set t" to achieve P_{FA}
4. Compute minimal failure size for given P_D.

Notice that, in contrast to Eq. 3-62, as we let N --> ∞, d^2 approaches a finite value. That is, there is a fundamental limit to the reliability of decisions due to the presence of model error -- an intuitively pleasing result.

EXAMPLE 2 (ISOLATION)

This example is a generalization of Example 1 and will have implications in terms of isolation test design for aircraft path failures. We consider fixed sample size tests for the problem of distinguishing the two hypotheses:

$$H_1 : y_k = m_1 + n_k \tag{3-80}$$

$$H_2 : y_k = m_2 + n_k \tag{3-81}$$

The truth model is described by

$$H_1 : y_k = m_1 + n_k + b \tag{3-82}$$

$$H_2 : y_k = m_2 + n_k + b \tag{3-83}$$

where m_1 and m_2 are two constant vectors and n_k and b are the white noise and random bias vectors of Example 1. The LLR test obtained using the design model Eqs. 3-80 and 3-81 is

$$\sum_{j=1}^{N} (m_1-m_2)^T \Sigma_n^{-1} y_j - \frac{1}{2} m_1^T \Sigma_n^{-1} m_1 + \frac{1}{2} m_2^T \Sigma_n^{-1} m_2 \underset{D_2}{\overset{D_1}{\gtrless}} t \tag{3-84}$$

As in Example 1 Eq. 3-84 can be rewritten as,

$$S = P^t \sum_{j=1}^{N} y_j \underset{D_2}{\overset{D_1}{\gtrless}} t' \tag{3-85}$$

The distinguishability metric, d^2, calculated using the Gaussian statistics of the truth model in Eqs. 3-82 and 3-83 is

$$d^2 = (N\, P^t \Delta m)^2 / \sigma^2 \tag{3-86}$$

where σ^2 is defined in Eq. 3-70 and $\Delta m = (m_1 - m_2)$. Given m_1 and m_2, the choice of P which maximizes d^2 is

$$P = \overline{C}^{-1}\, \Delta m \tag{3-87}$$

The choice of t' can be made for either H_1 or H_2 in a manner similar to Example 1. Alternatively, a Bayesian cost can be defined as in Eq. 3-50 and the decision regions computed by solving the Baysian optimization problem using the truth model statistics defined by Eqs. 3-82 and 3-83.

The similarity to Example 1 breaks down when we try to define minimal failures. This is because, for the test described by Eq. 3-85 with the projection vector defined by Eq. 3-87, the d^2 metric (Eq. 3-86) is not necessarily monotonic with failure size. That is, if the d^2 metric is computed using $c_i m_i$ in place of m_i in Eq. 3-86, then as c_i increases from 1, d^2 may go up or down. In fact, it can be shown in some cases that the probability of making an incorrect decision may actually approach 1 as c_i approaches infinity. This fact can easily be seen in Fig. 3-10. The figure shows the decision region in the space spanned by the decision statistics $\sum_k y_i(k)$ for i = 1, 2 (the components of a two dimensional measurement y(k). These decision regions are

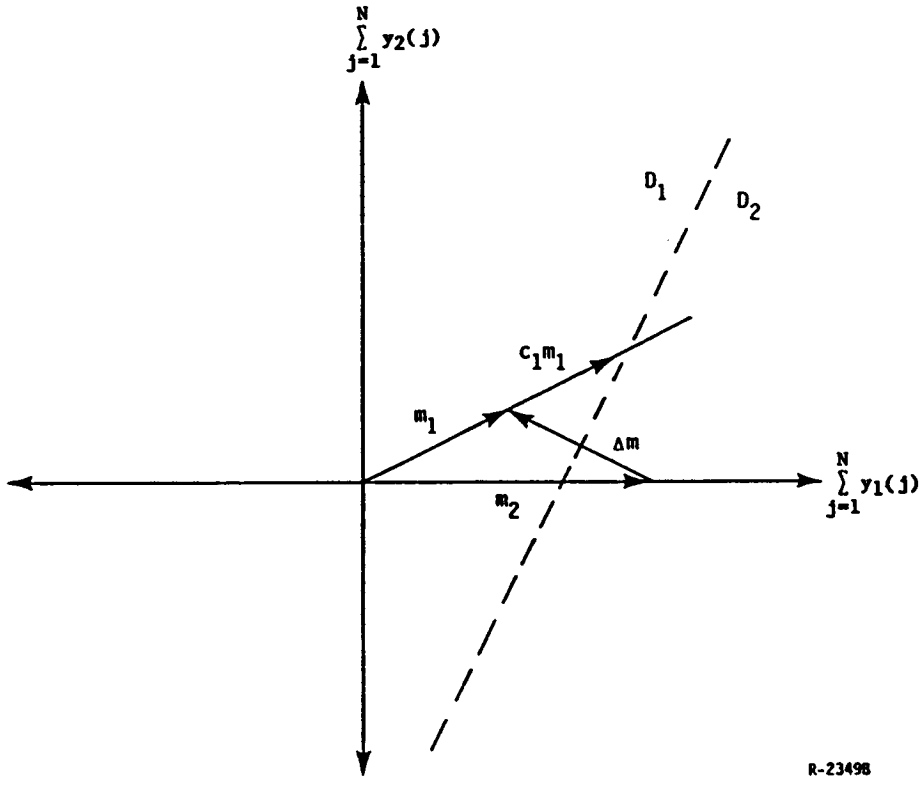


Figure 3-10. Failure Geometry in Measurement Space

determined by Eq. 3-85 with a threshold that weights the costs of incorrect decisions (for both $c_i = 1$) equally in a Baysian problem. It is easily seen from the figure that if the magnitude of m_1 is actually larger than its design value, then the likelihood that the statistic lies in D_2 can become quite large. Thus, this particular choice of parameters is <u>not</u> robust to changes in the magnitudes of the vectors m_i.

EXAMPLE 3 (ROBUST ISOLATION)

The problem in the previous example illustrates that we can not always use the structure defined by a design model and expect adequate performance. This is because some unknown parameters, such as failure size, result in variations in performance which are too large to be acceptable. When this is the

case, there are typically two approaches which can be taken. The first and most often used approach involves estimating the unknown parameters and forming a generalized likelihood ratio test [32]. While this approach performs well in some cases, it is also possible that, due to mismatch between the truth model and the real world, estimation errors can cause severe performance degradation. Furthermore, the analysis and design of such algorithms is complicated by the fact that a more detailed truth model is needed. The alternate approach, which is described here by example, is sometimes referred to as invariance. In this approach we try to develop tests whose performance is invariant to changes in these "nuisance" parameters.

In this case we consider the design model hypotheses,

$$H_i : y_k = c_i m_i + n_k \quad i = 1, 2 \tag{3-88}$$

where c_i is an unknown scalar. When both c_i were known, the binary decision problem of Example 2 completely specified both test structure and thresholds, and performance measures could be computed. In Example 1, however, notice that one option we had in the design process was to turn the problem into one of rejecting H_0 with high confidence (satisfy P_{FA}) and then <u>evaluate</u> what alternate hypothesis parameters (failure size) would ensure adequate performance (P_D). We take the same approach in this example for <u>each</u> hypothesis described by Eq. 3-88. We will see that c_i invariant decision regions using this approach can easily be described.

We start with the problem of rejecting H_1 using an N-window of observations. Using the results of subsection 3.3.1, we can define for any value of c_1, an α significance rejection contour by

$$D(c_1) = \{Y_k : \Pr[\| \sum_{k=1}^{N} (y(k) - c_1 m_1) \|^2 \geq \| \sum_{k=1}^{N} (Y_k - c_1 m_1) \|^2] \leq \alpha\} \tag{3-89}$$

Note y(k) denotes the random variable and Y_k are sample values. The sums in Eq. 3-89 are taken over the N-window of data and the norms are taken with respect to the white noise covariance matrix to be consistent with the design model Eq. 3-88. Since c_1 is unknown, however, we want a decision region for rejecting H_1 such that the significance of the resulting test is invariant under changes in c_1. That is, we want a decision region which is the locus of points in measurement space defined by Eq. 3-89 for all c_1, or

$$D = D(c_1) \tag{3-90}$$

For two dimensional measurements, y_k, this region is easily derived from Fig. 3-11. To define this region in general, let W represent a transformation matrix such that $z_k = W\, y(k)$ (note we can consider each time instant separately in this case because the mean vector m_1 is constant). If $W\, m_1 = 0$, then $p(z_k \mid H_1) = p(W\, n_k \mid H_1)$ which is not a function of c_1. If the measurement vector is n-dimensional, then the transformation matrix W spans an n-1 dimensional subspace. The generic rejection test would then be defined by the region D which can be rewritten,

$$D = \{\, Y_k : Z_k = W\, Y_k \text{ and } \Pr\Big[\sum_k (z_k{}^t C_z^{-1} z_k) > \sum_k Z_k{}^t C_z^{-1} Z_k) \Big] < \alpha \,\} \tag{3-91}$$

where C_z is some positive symetric matrix (usually the covariance matrix of Z_k, though others may be desirable).

Given a choice for C_z and Δc, the region D completely defines a rejection test for H_1 which is insensitive to changes in c_1. In this example, however, there is an alternate hypothesis which we have not yet discussed. Since there is only one alternate hypothesis, H_2, in this example, we would <u>like</u> to define

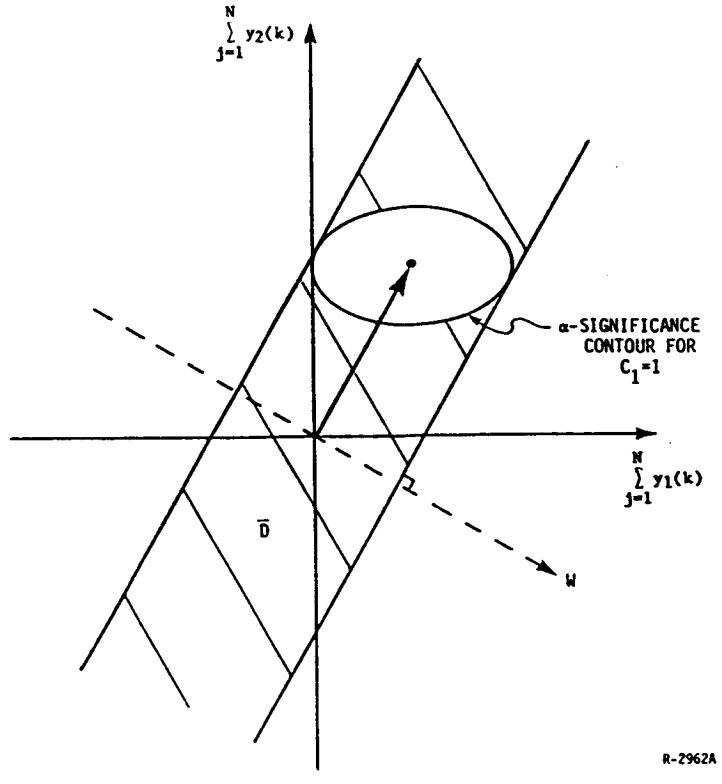


Figure 3-11. Two Dimensional Visualization of Decision Region $\bar{D}$ = Complement of D

the H_1 rejection region to be maximally sensitive to H_2 (rather than being generic). That is, we want to define D so that the smallest possible values of c_2 will make the probability that Y_k lies in D large. To do this, let us first define the rejection test based on the design model as:

$$S_1 = P_1{}^t \sum Y_k > D_1 \; t_1 \qquad (3\text{-}92)$$

Note that this test is not in the same form as Eq. 3-91 indicates, but is consistent with the test structures in Example 2. Again, this test assumes the failure sign is known. To be a rejection test for H_1 we know, from the above, that P_1 must be orthogonal to m_1. The remaining degrees of freedom are then

chosen to optimize the sensitivity to H_2 using, for example, the truth model of Eqs. 3-82 and 3-83. The d^2 metric is used now as a measure of sensitivity because, for the test defined by Eq. 3-92 with $P_1{}^t m_1 = 0$, d^2 is monotonic c_2. That is,

$$d_1{}^2 = [E\{S_1 \mid H_2\} - E\{S_1 \mid H_1\}]^2 \,/\, \mathrm{Var}[S_1 \mid H_1 \text{ or } H_2]$$
$$= c_2{}^2[P_1{}^t m_2]^2 \,/\, P_1{}^t[N^2\Sigma_b + N\Sigma_n]P_1 \qquad (3\text{-}93)$$

Equation 3-93 is maximized for any value of c_2 by choosing P_1 as follows. Let $\Sigma = N^2\Sigma_b + N\Sigma_n = Q^tQ$, $\bar{P}_1 = QP_1$, and $\bar{m}_i = Q^{-t}m_i$ for $i = 1, 2$. Then the desired optimization problem is,

$$\max \; [\, \bar{P}_1{}^t \bar{m}_2]/\bar{P}_1{}^t \bar{P}_1$$
$$\text{s.t. } \bar{P}_1{}^t \bar{m}_1 = 0 \qquad (3\text{-}94)$$

The objective function in Eq. 3-94 is easily seen as the projection of $\bar{m}_2$ onto P_1 and the constraint requires that $\bar{P}_1$ lie in the null space of $\bar{m}_1$. Thus the solution is just the projection of $\bar{m}_2$ onto the null space of $\bar{m}_1$. Finally solving for P_1, we have,

$$P_1 = K\, \Sigma^{-1}[m_2 - (m_2{}^t\Sigma^{-1}m_1)/(m_1{}^t\Sigma^{-1}m_1)m_1] \qquad (3\text{-}95)$$

where K is just a normalization constant.

We can now follow the same procedure for H_2 and obtain a test based on a decision statistic S_2, which rejects H_2 with maximal sensitivity to H_1. The two hypothesis tests define decision regions in which it is possible to reliably reject each hypothesis. This is shown in the two dimensional case in Fig. 3-12. Note that in contrast to Example 2, there are four regions of interest:

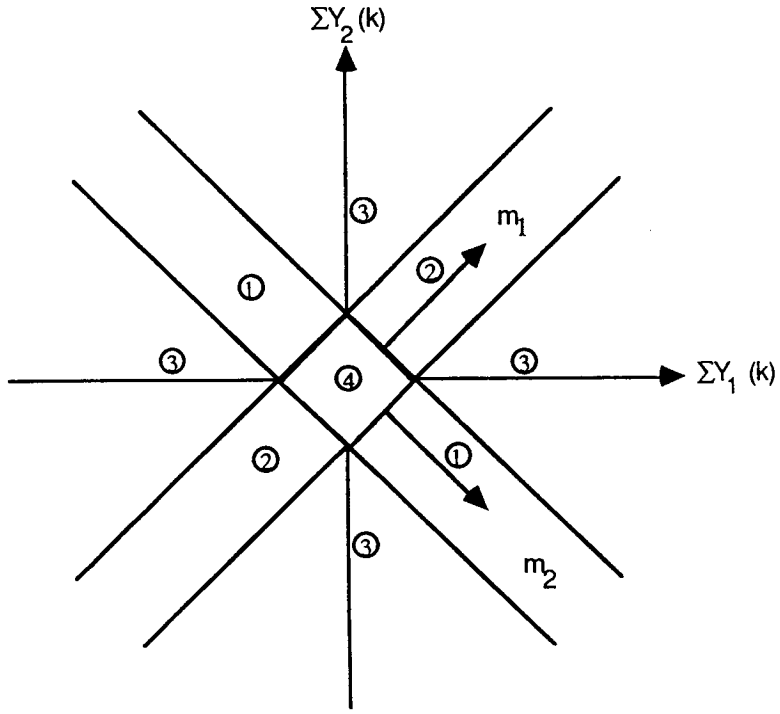


Figure 3-12. Decision Regions for Two Rejection Tests

Region 1) H_1 can be rejected, but H_2 can't ——> Decide H_2

Region 2) H_2 can be rejected, but H_1 can't ——> Decide H_1

Region 3) Both H_1 and H_2 can be rejected ——> Reject H_1 and H_2

Region 4) Neither H_1 nor H_2 can be rejected ——> Can't decide

Thus, these two tests allow us more flexibility in the decisionmaking process when events occur that can be described by neither of the hypotheses.

Finally, we note that it would be possible to simplify this two-test decision process by considering the test,

$$\begin{aligned} (S_1 - S_2) > t_1' &\longrightarrow \text{Decide H2} \\ (S_1 - S_2) < -t_2' &\longrightarrow \text{Decide } H_1 \\ \text{Otherwise} &\longrightarrow \text{Can't decide} \end{aligned} \qquad (3\text{-}96)$$

This test is not equivalent to the two test procedure. The can't decide mode may incorporate parts of both regions 3 and 4 of the two-test procedure depending on the choice of t_1' and t_2'. Clearly $t_1' = t_2' = 0$ is not a reasonable choice. However, the choice of t_1 and t_2 and the relationship of these thesholds to the two-test procedure thresholds has not been fully investigated. A reasonable choice would be to choose $t_1 = t_2$ such that, independent of the hypotheses failure size (c_i), the probability of making a wrong decision is guaranteed to be low. For example, if we want $Pr[(S_1 - S_2) > t \mid H_1) < \alpha$ independent of c_1, we assume $c_1 = 0$ (the worst case) and set the threshold at the α significance level for the assumed Gaussian distribution of $(S_1 - S_2)$.

EXAMPLE 4 (UNKNOWN FAILURE SIGNS)

In the above examples, we assumed that even if the "failure" magnitudes were unknown (c_i in Example 3), the sign of the failure was known. This example justifies a rather obvious modification to the tests described in the previous examples to account for uncertain signs. The modification results in taking the absolute value of the decision statistics, i.e., performing two-sided tests.

Consider the design model of Example 1 with uncertain failure sign given by,

$$H_o : y_k = n_k$$
$$H_1 : y_k = im + n_k \qquad (3\text{-}97)$$

where i takes on a value of 1 or -1. The generalized likelihood ratio test for an N-window of measurements Y^N is defined by,

$$\max_i p(Y_N \mid H_1, i) \;/\; \max_i p(Y_N \mid H_o, i) \underset{D_o}{\overset{D_1}{\gtrless}} t \qquad (3\text{-}98)$$

The denominator of Eq. 3-98 is not a function of i and the numerator can be maximized by maximizing the natural log of the numerator (assume Gaussian statistics and n_k white as in Example 1). Ignoring the terms which don't depend on the choice of i, we get

$$i = \arg \max_i \left[m \sum_{k=1}^{N} y_k \right] \qquad (3\text{-}99)$$

If m and $\sum_{k=1}^{N} y_k$ are the same sign, i = 1; and if they are opposite signs, then i = -1. Thus, $i = \text{sgn}\left(m \sum_{k=1}^{N} y_k\right)$. Putting this solution back into Eq. 3-98, taking logarithms, incorporating all the constants into a threshold and expressing the decision statistic in general form (as we did in Example 1) we have the test,

$$\left| P^t \sum_{k=1}^{N} y_k \right| \underset{D_o}{\overset{D_1}{\gtrless}} t' \qquad (3\text{-}100)$$

SEQUENTIAL TEST DESIGN

The design of sequential tests such as the SPRT mentioned in subsection 3.3.1 is somewhat different than the fixed sample size tests described in the above examples. Sequential tests address the binary decision problem only. Samples (measurements) are taken until a decision can be made in favor of one hypothesis or the other. The number of samples needed to complete the test is a random variable. Characteristics of both hypotheses must be specified in some detail which, for the FDI problem implies that we must choose "minimal" failure magnitudes and ensure that larger failures result in shorter test times.

As discussed at the very beginning of this subsection, we want to specify sequential tests using a design model and then choose their parameters based on a truth model. Consider as an example, the design and truth models of Example 1. The SPRT is easily derived and written in a general form as

$$S_k = G_I \sum_{j=1}^{k} (P^t y_j) - \sum_{j=1}^{k} G_D \; > \; t^+ \longrightarrow \text{Decide } H_1$$

$$< \; t^- \longrightarrow \text{Decide } H_0$$

$$< t^+ \text{ and } > t^- \longrightarrow \text{Take another sample} \qquad (3\text{-}101)$$

where G_I, G_D, P, t^+, and t^- are parameters to be specified. Furthermore, since we know that model error exists, the theoretical guarantee that this procedure terminates is no longer valid. Thus, it is often useful to specify a time limit. When the time limit is reached and no decision has been made, we may either exit with no-decision as a conclusion or perform a fixed sample size test at this time. A prototype design procedure which makes use of the truth model is given below.

Step 1:

Choose a maximum sample length, N.

Step 2:

Select P to maximize the distinguishability metric, $d^2(N)$, for the maximal length test as in Example 1 (the answer is the same since d^2 is not a function of G_I or G_D)

Step 3:

In keeping with the ideal parameter values $G_I\ p^t = m^t\Sigma_n$ and $G_D = m^t\Sigma_n m/2$ we proceed as follows. Determine the value of $\bar{S} = E\{\sum_{k=1}^{N} p^t y_k\}$ which makes $d^2(N)$ acceptable. Set $G_1 = \bar{S}$ and $G_D = \bar{S}^2/2$. This results in a test in which values of $\sum_j p^t y_j$ less than $\bar{S}/2$ tend to drive the statistic to its negative threshold.

Step 4:

Determine thresholds $t^+ = -t^-$.

The last step is the most difficult since no closed form solution exists which relates the desired performance measures (e.g., P_{FA} and P_D) to the choice of thresholds when the truth model and the design model are not identical. Reference [35] presents a numerical method which could be used for such calculations, however, the results are incorrect as detailed in [36]. The difficulty is easily illustrated by the following equations.

Let S_k be the decision statistic at stage k in the SPRT defined in Eq. 3-101 with thresholds t^+ and t^-. Also, let $S^i = (S_1, S_2, \ldots, S_i)$. We can relate P_{FA} and P_D to t^+ and t^- by,

$$P_{FA} = \sum_{i=1}^{N} \int_{L^+} P(S^i \mid H_o)\, dS^i$$

$$L^+ \triangleq \{S^i : t^- < S_j < t^+,\ j=1, \ldots, i-1 \text{ and } S_i > t^+\} \quad (3\text{-}102)$$

$$P_D = \sum_{i=1}^{N} \int_{L^+} P(S^i \mid H_1)\, dS^i \quad (3\text{-}103)$$

Also, we have probabilities of correct rejection (P_{CR}) and missed detection (P_{MD}) which are given by,

$$P_{CR} = \sum_{i=1}^{N} \int_{L^- \triangleq \{S^i : t^- < S_j < t^+,\ j=1, \ldots, i-1 \text{ and } S_i < t^-\}} P(S^i \mid H_o)\, dS^i \qquad (3\text{-}104)$$

$$P_{MD} = \sum_{i=1}^{N} \int_{L^-} P(S^i \mid H_1)\, dS^i \qquad (3\text{-}105)$$

Note that for $N < \infty$, we have $P_{FA} + P_{CR} \neq 1$ since there is a finite probability that no decision is made (i.e., $t^- < S_j < t^+$, $j=1, \ldots, N$). Also, Eqs. 3-102 - 3-105 assume that no decision is made if $S^N \in L^+$ or L^-. In some cases, it may be appropriate to define the terminal decision rule $S_N \underset{D_o}{\overset{D_1}{\gtrless}} \bar{t}$ if $S^N \in L^+$ or L^-. In this case Eqs. 3-102 - 3-105 would have to be modified.

Equations 3-102 through 3-105 must now be evaluated using the truth pdfs of S_k (as opposed to those assumed in defining S_k as a function of y_k). Unfortunately, closed form solutions for these equations are available only in the simplest cases. If, for example, S_k was an uncorrelated Gaussian process (not likely, in view of equations like Eq. 3-101), then we could write

$$P_{FA} = \sum_{i=1}^{N} p_i^{\circ} \prod_{m=1}^{i-1} (1-r_m^{\circ}) \qquad (3\text{-}106)$$

$$P_{MD} = \sum_{i=1}^{N} q_i^{\circ} \prod_{m=1}^{i-1} (1-r_m^{1}) \qquad (3\text{-}107)$$

$$P^{j}_{\text{No Decision}} = \prod_{i=1}^{N} (1-r_i^{j}) \qquad (3\text{-}108)$$

where $p_i^j \triangleq \Pr\{S_i > t^+ | H_j\}$ = Probability of Deciding H_1 at time i, given H_j and no decision up until time i

$q_i^j \triangleq \Pr\{S_i < t^- | H_j\}$ = Probability of Deciding H_0 at time i, given H_j and no decision up until time i

$r_i^j \triangleq 1-p_i^j - q_i^j$ = Probability of deciding to take another sample at time i, given H_j and no decision up until time i.

Note that p_i^j and q_i^j are easily computed if S_i is Gaussian.

While Eqs. 3-106 - 3-108 may provide approximations to the desired quantities when S_i is not uncorrelated, it is not clear how accurate the approximation is or when it would break down. In order to get more accurate estimates of the desired quantities, numerical integration or simulation methods are needed. For the control element FDI algorithm developed in Section 5, some interesting heuristic methods of computing thresholds for sequential tests are developed.

THRESHOLD SCHEDULING

The computations described above are useful when errors in the pdf specifications can only be described in a statistical sense. If this statistical description adequately characterizes the worst errors throughout the operational region (or envelope) of the system under consideration, we can be quite

confident that the resulting FDI design will meet its performance specifications under all conditions.

The problem in such a worst-case analysis, however, is that the errors encountered at the limits of this envelope may be so severe that exceptionally conservative designs result. That is, only failures of extreme severity would be large enough to appear significant over the worst-case modeling error.

The natural alternative to a worst-case design is one in which we consider a limited envelope for our baseline design and modify the algorithm towards a more conservative design when large errors are anticipated. The mechanism which accomplishes this adjustment is known as "threshold scheduling" and has proven to be an important part of many FDI systems [5], [14] - [16]. The basic idea behind threshold scheduling can be illustrated by considering the problem of detecting inertial sensor failures using dual redundant sensors while accounting for misalignment errors. The residual, formed by comparing two similar sensors at each time, k, in this case is given by

$$\nu(k) = y_1(k) - y_2(k) \tag{3-109}$$

where

$$y_i(k) = a_i x(k) + \eta_i(k)$$

$x(k)$ = variable being measured

a_i = actual scale factor (both equal to 1 with zero alignment error)

$\eta_i(k)$ = zero mean white Gaussian noise

Thus, we can express the residual, under normal operation, as,

$$\nu(k) = \Delta a\, x(k) + N(k) \tag{3-110}$$

where $\Delta a = a_1 - a_2$ and $N(k) = \eta_1 - \eta_2$. Now, in order to achieve a false alarm rate of 1 in 10,000, we know that (using percentage points for the normal distribution) our thresholds (if we perform no information collection) must be $\pm 3.5\ \sigma_N$ when either $\Delta a = 0$ or $x = 0$. The problem is that large values of x result in large residuals when Δa is non-zero, and the term $\Delta a x(k)$ may get, in the worse case, sufficiently large so that only enormous differences between y_1 and y_2 can be detected with a threshold that guarantees P_{FA} for worst-case errors. Threshold scheduling overcomes this worst-case performance by adjusting the threshold on the basis of a failure-insensitive, deterministic estimate of x(k) and a specification of the worst-case scale factor differential. Thus, the sensor FDI test becomes,

$$| \nu(k) | > 3.5\sigma_N + \Delta a_{max} \hat{x}(k) \text{ ===> Failure Detected} \qquad (3\text{-}111)$$

The above example illustrates the need for threshold scheduling and the basic concepts involved (note, that no attempt is made here to elaborate on how a failure-insensitive estimate could be obtained). In order to apply these ideas to other FDI designs, we now describe some of the general principles involved.

First, we note that the errors which are important for threshold scheduling are those whose statistical contribution to error in the residual depends on a deterministic signal. In the example above, we saw that the deterministic signal $\hat{x}$ was used to schedule thresholds. Such a signal provides knowledge about the potential for large errors at each point in time.

Next, the statistical nature of the error source is described by specifying the pdf of the error as a function of this deterministic signal. Again, in the above example, we effectively assumed that Δa was uniformly distributed

on $[-\Delta a_{max}, \Delta a_{max}]$ resulting in an error term which is uniformly distributed on $[-\Delta a_{max}\hat{x}(k), \Delta a_{max}\hat{x}(k)]$. Note, we could also have assumed that Δa was Gaussian with a known variance, and obtained a different scheduling algorithm.

The threshold schedule can now be determined by specifying the performance characteristic which must be preserved throughout the entire operational envelope. For example, the threshold in Eq. 3-111 is the result of requiring that the false alarm performance be maintained. If missed detection performance were to be maintained, we would need to characterize the pdf of the residual under failure conditions along with the error and compute the scheduled thresholds accordingly. Note that we cannot maintain both performance measures simultaneously in a fixed sample test since only a single threshold can be modified.

Sequential Test Scheduling - In sequential testing procedures such as the SPRT, two threshold must be scheduled. Since, in these tests, we are willing to declare that no decision can be made, it is, therefore, possible to maintain both P_D and P_{FA}, although as errors get large, the likelihood of not making any decision increases. As with the determination of nominal thresholds, the computations involved in determining threshold schedules for sequential tests are more complex than the fixed sample size case. However, a reasonable approach is to choose the threshold at each stage, k, such that the probabilities $P(S_k>t^+|H_0)$ and $P(S_k<t^-|H_1)$ (conditions which, respectively, result in a false alarm and a missed detection at stage k) are equivalent to those achieved in the original SPRT with no error. (Note, S_k is the SPRT decision statistic as, for example, in Eq. 3-101). Such a procedure is simple since each threshold only depends on an estimate of the impact of errors on S_k

independent of other estimates at other stages, and has performed quite well in practice [5].

To see this, consider Example 1 with the SPRT defined by Eq. 3-101. Suppose that in addition to the random bias error in Eqs. 3-68 and 3-69, there is an additional error term E_k, which is a function of a known deterministic signal that is independent of which hypothesis is true. That is, the truth model is now

$$\begin{aligned} H_o \quad & y_k = \eta_k + E_k + b \\ H_1 \quad & y_k = \eta_k + E_k + b + m \end{aligned} \tag{3-112}$$

Let t_o denote the threshold we select assuming $E_k = 0$, for all k, and let $t_E(k)$ denote the threshold at time k based on the criterion that the probability of false alarm at stage k, given that we've reached stage k, is the same for $E_k = 0$ and $E_k \neq 0$. That is we want to ensure that $Pr[S_k > t_o \mid H_o, E = 0] = Pr[S_k > t_E(k) \mid H_o, E \neq 0]$. Since S_k is Gaussian, this condition implies that:

$$t_o - \bar{S}_{oo}(k) = t_E(k) - \bar{S}_{OE}(k) \tag{3-113}$$

where $\bar{S}_{oo}(k) = E\{S_k \mid H_o, E = 0\}$, and $\bar{S}_{OE}(k) = E\{S_k \mid H_o, E \neq 0\}$. For this example, we have,

$$t_E(k) = t_o + \sum_{j=1}^{k} G_I P^t E_j \tag{3-114}$$

Equation 3-114 defines a threshold scheduling algorithm for the SPRT of Eq. 3-101. If the sign of E_k is unknown, we would like to use the worst case combinations and so the second term in Eq. 3-114 becomes

$$\sum_{j=1}^{k} G_I P^t \, \mathrm{Sgn}(E_j P^t) \, E_j \qquad (3\text{-}115)$$

where $\mathrm{Sgn}(E_j P^t)$ is a diagonal matrix of 1s and -1s such that if $[E_j p^t]_{ii} > 0$, then $[\mathrm{Sgn}(E_j p^t)]_{ii} = 1$ and if $E_j P^t < 0$ then $[\mathrm{Sgn}(E_j p^t)]_{ii} = -1$. For implementation purposes we note that $P_t \, \mathrm{Sgn}(E_j P^t) E_j = \sum_{i=1}^{n} | P_i E_j{}^i |$ where n is the dimension of P and E_j and P_i and $E_j{}^i$ denote the i^{th} element of P and E_j respectively.

Scheduling in Single Input Single Output (SISO) Systems with Transfer Function Errors - The The problem to be formulated here is motivated by the problem of detecting "actuator path" failures as shown in Fig. 3-13. In this problem, we have nearly perfect measurements of both the input to the actuator and its output.

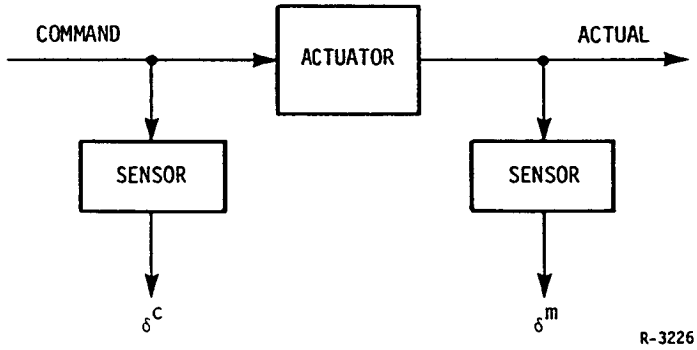


Figure 3-13. Measurements for Actuator Path Failures

Figure 3-14 shows the "open loop" residual generation process for linear single-input, single output (SISO) systems (see subsection 3.2 for definition of open-loop residuals). Although closed loop or finite memory residual processes may be considered, we consider only the open loop case since it has been used successfully in cases where the SISO system is a stable, high bandwidth low-pass system.

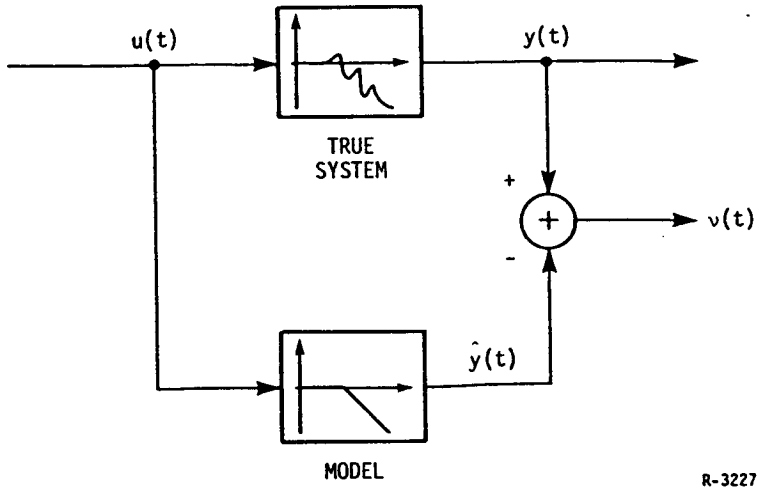


Figure 3-14. Open-Loop Residual Generation in SISO Systems

The residual process, $\nu(t)$ in Fig. 3-14, is non-zero because the true transfer function from u to y is different than the model we use to produce $\hat{y}$. Typical model-errors are due to high frequency dynamics and nonlinearities which cannot be easily characterized and variations in hardware (over time and from implementation to implementation). The residual process $\nu(t)$ is a non-stationary process because the input, u(t) is nonstationary. Furthermore, the size of the residual scales with the size of u(t). To see this, consider the case where both the true system and the model are linear systems. In this case, we have

$$\nu(t) = \int_0^t e(\tau)\, u(t-\tau)\, d\tau \qquad (3\text{-}116)$$

where,

$$e(\tau) = h(\tau) - \hat{h}(\tau)$$

$h(\tau)$ = true-system impulse response

$\hat{h}(\tau)$ = model

To design an FDI system which detects failures of the true system, we must first characterize the residual process when no failure exists. Since, in Eq. 3-116, $e(\tau)$ is not known, and in fact may vary over time and implementations of $h(t)$, let us characterize it by the functional relationship $e(\tau;\theta)$ where θ is an unknown random paramter. The parameter, θ, is the only random element in the description of $\nu(t)$. Although $u(t)$ could be characterized in some cases by a piecewise-stationary stochastic process, such a characterization would be of little use since we have a perfect measurement of u.

The "size" of $\nu(t)$ with no failure present can now be characterized as a function of the complete past history of the input; i.e., $\{u(\tau),\ \tau\varepsilon\ [0,t]\}$. To see this, let us compute the mean square value of $\nu(t)$ with respect to the variations in θ. That is, if we compute;

$$\overline{\nu^2} = \mathcal{E}_\theta\{\nu(t)^2\}$$

where $\mathcal{E}_\theta$ denotes expectation with respect to the distribution of θ, then $\overline{\nu^2}$ can serve as a measure of the possible size of the residual under no failure and used as a basis for FDI.

To compute $\overline{\nu^2}$, we formally have,

$$\overline{\nu^2} = \mathcal{E}_\theta \left\{ \left[\int_0^t e(\tau;\theta)\, u(t-\tau)\, d\tau \right]^2 \right\} \tag{3-118}$$

$$= \left\{ \int_0^t \int_0^t d\tau_1\, d\tau_2\, f(\tau_1,\tau_2)\, u(t-\tau_1)\, u(t-\tau_2) \right\} \tag{3-119}$$

where

$$f(\tau_1,\tau_2) = \mathcal{E}_\theta \{e(\tau_1;\theta)\, e(\tau_2;\theta)\} \qquad (3\text{-}120)$$

Thus, we see that knowledge of u over the interval [0,t] and a statistical characterization of the error impulse response function is sufficient for computing $\overline{\nu^2}$.

Unfortunately, the calculation in Eq. 3-119 is not an efficient one since, in general, $f(\tau_1,\tau_2)$ is not seperable (if $f(\tau_1,\tau_2) = f_1(\tau_1 \mu\sim 2[\tau_2)$ then Eq. 3-119 amounts to the product of two linear filters). To see this, let's consider a simple example (of no particular interest to the actuator FDI problem).

Suppose we let $e(\tau;\theta)$ be given by,

$$e(\tau;\theta) = \theta e^{-\theta\tau} \qquad (3\text{-}121)$$

That is, the error transfer function in a first order low pass filter with unknown cut-off frequency, θ. This might correspond to a case where the true system is well known at high frequencies but uncertain at low frequencies. Note that Eq. 3-121 specifies a magnitude and phase relationship between u and ν. Many characterizations of unmodelled dynamics presume only a magnitude relationship with unknown phase [37]. This characterization will be discussed subsequently.

If we denote the probability density function of θ by $p(\theta)$, then;

$$f(\tau_1,\tau_2) = \int p(\theta)\, \theta^2 e^{-\theta(\tau_1+\tau_2)}\, d\theta \qquad (3\text{-}122)$$

If, in addition, we assume that θ is uniformly distributed over the interval $[\theta_L, \theta_H]$, then, using integration by parts

$$f(\tau_1,\tau_2) = \int_{\theta_L}^{\theta_H} \frac{1}{\theta_H-\theta_L}\ \theta^2 e^{-\theta(\tau_1+\tau_2)}\ d\theta \qquad (3\text{-}123)$$

$$= \frac{1}{\theta_H-\theta_L}\left\{ \frac{\theta^2 e^{-\theta(\tau_1+\tau_2)}}{-(\tau_1+\tau_2)} - \frac{2}{(\tau_1+\tau_2)^3}\ e^{-\theta(\tau_1+\tau_2)}\ (1+\theta(\tau_1+\tau_2)) \right\} \Bigg|_{\theta_L}^{\theta_H} \qquad (3\text{-}124)$$

There is no apparent way to separate Eq. 3-124 into a product of functions, each involving only τ_1 or τ_2. In addition, even this example demonstrates the complexity of calculating $\overline{\nu^2}$, using Eq. 3-119.

Part of the complexity in the above formulation is due to the significant amount of structure imposed on the error transfer function. As we mentioned above, transfer function errors are more commonly specified in terms of their magnitude response alone, with no knowledge available about phase. This leads us to search for scheduling methods which are based on the magnitude (or squared magnitude) of the Fourier transform of $e(\tau)$, viz. $E^2(\omega)$. We can do this using Parseval's relation as follows. Define a positive function of frequency $L(\omega)$ such that $L^2(\omega) \geq E^2(\omega)$ for all ω is guaranteed. Let $N(\omega)$ represent the Fourier transform of the residual sequence $\nu(t)$ and $U(\omega)$ be the Fourier transform of the input $u(t)$. Then the following relations hold:

$$N^2(\omega) = E^2(\omega)\ U^2(\omega) \qquad (3\text{-}125)$$

$$N^2(\omega) \leq L^2(\omega)\ U^2(\omega) \qquad (3\text{-}126)$$

$$\int_{-\infty}^{\infty} N^2(\omega)\ d\omega \leq \int_{-\infty}^{\infty} L^2(\omega)\ U^2(\omega)\ d\omega \qquad (3\text{-}127)$$

and using Parsevals relation,

$$\int_{-\infty}^{\infty} \nu^2(t)\, dt \leq \int_{-\infty}^{\infty} \left[\int_{-\infty}^{t} |\ell(\tau)\, u(t-\tau)\, d\tau \right]^2 dt \qquad (3\text{-}128)$$

where $\ell(t)$ is the inverse Fourier transform of $L(\omega)$. Equation 3-128 suggests a test which will reliably reject the hypothesis that a SISO system with transfer function error (i.e., error between reality and the model used to form a residual), is operating normally. The right hand side of Eq. 3-128 says that the input is filtered using any filter whose squared magnitude in the frequency domain bounds the actual frequency domain error. The output of this filter is then squared and integrated. This is then compared to the integral of the squared residual. The right hand side of Eq. 3-128 is the threshold and it is scheduled dynamically based on the temporal characteristics of the input. Of course, in a practical system, we would deal with these relationships in the discrete time domain and perform sums only over finite intervals to ensure that the values in Eq. 3-128 remain bounded.

OTHER DESIGN ISSUES

The design of fixed sample size hypothesis tests discussed in this subsection have all been based on a steady state analysis of the decision statistics which are used for decisionmaking. That is, only the steady state pdfs are used to optimize parameters and compute thresholds. In using these tests in the trigger process (see subsection 3.1), we clearly would like to minimize the number of samples required or equivalently the bandwidth of the resulting filter. This is because we recognize that there is a tradeoff between the size of failures one can detect and the speed in which this detection can take place. Some new ideas which are based on the notion of transient gain are derived here.

The transient gain of a filter $H(\omega)$ (or its Fourier transform $h(t)$) is defined as

$$G_T(h) = \sup J_T(h * y) \; / \; J_T(y)$$

$$y \in L_2(0,T) \tag{3-129}$$

where * denotes convolution, and where

$$J_T(x) = 1/T \int_0^T |x(t)|^2 \, dt \tag{3-130}$$

In [38] $J_T(\nu)$, (where ν is the residual vector of a Kalman filter), is used as a trigger statistic and the "size" of J_T is established for various values of T using uniformly distributed modeling and sensor errors.

The notion of transient gain is one which may be useful in designing trigger filters. To simplify the results which follow, we will deal with discrete time processes and finite impulse response (FIR) filters. The discrete version of Eq. 3-129 is

$$G_T(H) = \sup \| H_T Y_T \| \; / \; \| Y_T \| \tag{3-131}$$

where H_T is the impulse response matrix, viz.

$$H_T = \begin{bmatrix} h_0 & 0 & 0 & \cdots & 0 \\ h_1 & h_0 & 0 & \cdots & 0 \\ h_2 & h_1 & h_0 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \vdots \\ h_T & h_{T-1} & & \cdots & h_0 \end{bmatrix} \tag{3-132}$$

and $Y_T^t = (y(0), y(1), \ldots, y(T))$. Now, in the design of trigger filters (choosing H) we would like to maximize the transient response to a failure subject to a constraint on the steady state variance of the output. That is, given the variance of the output of this filter, we can first choose a fixed threshold which achieves some desired false alarm rate. Of all the filters which result in the same output variance, we would like the one which responds as quickly as possible. Formally stated, we want to choose the (T + 1) impulse response coefficients which

$$\max G_T(h)$$

subject to:

$$\underline{h_T}^t C_y \underline{h_T} = \sigma^2 \qquad (3\text{-}133)$$

where $\underline{h_T}$ is the vector of impulse response coefficients (i.e., the first column of H_T) and C_y is the autocorrelation matrix for Y_T. The solution of this problem requires a gradient scheme since both the objective and constraints are nonlinear in the decision vector h_T. Such schemes would be facilitated by observing that,

$$G_T(H) = \lambda_{max} (H_T H_T^t) = \sigma_{max} (H_T) \qquad (3\text{-}134)$$

where λ_{max} denotes the maximum eigenvalue and σ_{max} denotes the maximum singular value of the corresponding matrix.

Other options which are similar to the above (some of which are easier to solve) are given below.

1. Minimize the variance of the output subject to a constraint transient gain.

2. Minimize the variance subject to a fixed step-response transient gain. That is, assume that y(j) is a constant in computing G_T.

This is solved by taking the filter coefficients as C_y^{-1} times a vector of ones. This is also the solution to the maximization of the ratio of step response to variance of the output.

3. Maximize the norm of the impulse response vector, h_T, subject to a constant variance of the output. This is equivalent to maximizing the magnitude of the output at time T subject to a unit norm input and a constant output variance. The solution to this problem is to take h_T as the eigenvector corresponding to the smallest eigenvalue of C_Y. Note that the similarity between this result and the robust parity check results of [24] are not coincidental since the problem statements are effectively equivalent.

4. Maximize the worst case transient gain (by selecting y(j) to minimize G_T) subject to a constraint on steady state variance. That is, select h_T to

$$\max \min G_T(h)$$

$$h_T \; y(j) \; : \; j=0, \ldots, T$$

subject to:

$$h_T{}^t \; C_y \; h_T = \sigma^2 \qquad (3\text{-}135)$$

3.4 SUMMARY

In this section we have developed a general structure for FDI systems and discussed many details associated with the design of each element of this structure. At the top level, the FDI function is decomposed into residual generation and decision processes. In the residual generation process, information about normal (unfailed) system redundancy (including temporal, direct as well as analytic redundancy) is used to form residual signals which are well behaved under no failure conditions and which deviate in easily characterized and distinguishable ways when failures occur. While Kalman filtering methods can be used to generate residuals without an explicit analysis of redundancy, the failure to assess redundancy and their associative errors can result in poor performance. This is because Kalman filtering is a centralized

approach to residual generation (using the system model as a whole) and will tend to mix well known relationships with poorly known relationships. The decentralized approach to residual generation breaks up the system model into individual relationships among measured variables (static and dynamic balance equations). These relationships can then be examined independently in terms of their usefulness in each part of the decision process.

The decision process collects information contained in the residuals over time by noninvertibly compressing the residuals into various decision statistics and comparing these statistics to threshold (or to each other in some cases). Because of the need to process residuals over time, the unknown failure onset time plays a large role in determining the structure of the decision process. Many decision mechanisms have been proposed for dealing with the uncertainty about failure onset. We have adopted the structure used in [14] which consists of trigger, verify, and isolate subprocesses. The trigger process is used as a quick response alarm to indicate the possibility of a failure. Its thresholds are set to ensure quick detection of important failures. Furthermore, the sensitivity to all important failures are maximized by designing a separate trigger test for each failure mode rather than using the common practice of mixing all information together in a rejection test for normal operation (e.g., WSSR). The verify process is used to achieve the false alarm rate specifications by setting its thresholds to reliably reject false triggers. The isolation process is run in parallel to the verify process and performs binary hypothesis tests comparing all pairs of triggered failures.

Finally, in designing the hypothesis tests which comprise the decision process, we defined a design procedure which involved four major steps. The

first step was to define an algorithm structure for each hypothesis test using a simplified "design model" which statistically describes the nominal behavior of the residuals under each hypothesis of interest. The second step was to select the parameters of this algorithm using averages taken with respect to a truth model. The truth model also statistically characterizes the residuals under the alternative hypotheses, however, it includes variations due to modeling error that was neglected in the design model. Although it would be conceivable to design an algorithm for the truth model in the first place, the point of using separate models for design and analysis is to gain confidence in algorithm robustness. Finally, sensitivity analyses are made to see how performance varies with respect to uncertainties which are not easily characterized by a statistical truth model (such as failure magnitude). Iterations between this process and the first and second step may be necessary.

SECTION 4

ANALYTIC METHODS OF EVALUATING FUNDAMENTAL LIMITS TO FDI PERFORMANCE

In the generic FDI design methodology discussed in subsection 3.1, we argued that the evaluation of fundamental limits to FDI performance was necessary both to determine feasibility of performing the FDI function and in determining sensitivities of performance to different system characteristics. In this section we develop an approach for performing these analyses based on the concept of distinguishability metrics. Such metrics were used in the design process as an objective function for the optimization of algorithm parameters. In this section we will assume that a statistical truth model of the observed quantities which adequately characterizes all important uncertainties is available for every hypothesis (mode of operation). For characteristics which can't be described statistically (e.g., failure size and signature), we will perform a parametric analysis of performance limitations.

4.1 DISTINGUISHABILITY METRICS

The basic idea behind our approach is to evaluate the distinguishability of pairs of hypotheses in terms of the smallest likelihood of making an incorrect decision based on some set of observations. This likelihood is nonzero because each hypothesis can only be described statistically.

It is well known [32] that the decision rule which achieves the smallest error probability, P_e, is the maximum a-posteriori (MAP) decision rule which achieves

$$P_e = \int \min \left[p(Z_p|H_i),\ p(Z_p|H_j)\right] dZ_p \qquad (4\text{-}1)$$

Equation 4-1 is, however, not particularly easy to evaluate even using numerical integration algorithms and even when the densities are Gaussian. For this reason, it is useful to define measures which can be related to P_e, but are easier to evaluate. Such measures and their application to FDI were the topic of references [22] - [24] and have found wide applicability in areas such as patten recognition [39], [40], control systems [41] - [43], communications [44], information theory [45], [46], and statistics [47] - [48]. We now define some of these measures and describe some of their useful properties.

If we view the pdf of a random variable, say ζ, as a vector in an infinite dimensional space, then the 'distance' between the two vectors (pdfs), $p_i(\zeta)$ and $p_j(\zeta)$, can be computed in a variety of ways. For example, the standard Euclidean or 2-norm is defined as

$$L_2 = \int \left[p_i(\zeta) - p_j(\zeta)\right]^2 d\zeta \qquad (4\text{-}2)$$

The Kolmogorov distance, or 1-norm, is defined as

$$K = \int \left|p_i(\zeta - p_j(\zeta)\right| d\zeta \qquad (4\text{-}3)$$

These two distance measures, while retaining some very nice topological properties [44], are either difficult to compute (in the case of K) or not easy to relate to P_e (in the case of L_2). In reference [24], two measures which can be easily computed and related to P_e were identified. They are the J-divergence which is defined by

$$J_{ij}(\zeta) = \int \left[p_i(\zeta) - p_j(\zeta)\right] \ln \frac{p_i(\zeta)}{p_j(\zeta)} d\zeta \qquad (4\text{-}4)$$

and the Bhattacharyya distance which is defined by

$$B_{ij}(\zeta) = -\ell n \int [p_i(\zeta)\, p_j(\zeta)]^{1/2} d\zeta \qquad (4\text{-}5)$$

Some general properties of J and B are given below.

1. $J_{ij} > 0$
2. $J_{ii} = 0$
3. $B_{ij} > 0$
4. $B_{ii} = 0$

Also note that neither J nor B necessarily satisfy the triangle inequality; i.e., for $i \neq j \neq k$,

$$J_{ij} + J_{jk} \neq J_{ik}$$

$$B_{ij} + B_{jk} \neq B_{ik}$$

The most important property of J and B is that they can be used to bound P_e as defined in Eq. 4-1. The following relationships are derived in [44] - [46]

$$1/2 - 1/2(1 - 4q_iq_j\, e^{-2B_{ij}})^{1/2} \leq P_e \leq (q_iq_j)^{1/2}\, e^{-B_{ij}} \qquad (4\text{-}6)$$

$$1/2\, \min(q_i,q_j)\, e^{-J_{ij}/8} \leq P_e \leq (q_iq_j)^{1/2}\, (J_{ij}/4)^{-1/4} \qquad (4\text{-}7)$$

with the upper bound in Eq. 4-7 valid only for Gaussian densities. In Eqs. 4-6 and 4-7, q_i and q_j are the a-priori probabilities associated with the two densities $p_i(\zeta)$ and $p_j(\zeta)$.

In the Gaussian case (which is of frequent interest) both metrics are relatively simple to compute. For example, if $p_i(\zeta) = N\{m_i;\ C_i\}$ and $p_j(\zeta) = N\{m_j;\ C_j\}$, then we have

$$B_{ij} = 1/2\ \ell n \det\{1/2\ (C_i+C_j)\ C_i^{-1/2}C_j^{-1/2}\} + 1/8\ (m_i-m_j)^T\left[\frac{C_i+C_j}{2}\right]^{-1}(m_i-m_j) \tag{4-8}$$

$$J_{ij} = 1/2\mathrm{Tr}\{C_jC_i^{-1} + C_iC_j^{-1} + [C_i^{-1} + C_j^{-1}]\ [m_i-m_j]\ [m_i-m_j]^T\} - t \tag{4-9}$$

where t is the dimension of ζ.

When the vector ζ represents a time series, some other interesting results can be derived. Let

$$\zeta = \{z(1),\ z(2),\ z(3),\ \ldots z(N)\} \tag{4-10}$$

where z(k) is a stationary Gaussian vector time series. For a particular hypothesis, say H_i, this time series is completely characterized by its (time varying) mean sequence, $m_j(k)$, and its covariance function, $C_j(k)$, where,

$$m_i(k) = E\ \{\ z(k)\ |\ H_i\ \} \tag{4-11}$$

$$C_i(k) = E\ \{(z(j) - m_i(j))\ (z(j+k)\ m_i(j+k))^t\ |\ H_i\} \tag{4-12}$$

The distinguishability metrics discussed above can be computed for the vector, ζ, from its mean vector $m_\zeta{}^i$, and (Toeplitz) autocovariance functionmatrix $C_\zeta{}^i$, which are functions of m_i and C_i in Eqs. 4-11 and 4-12, respectively. Thus, given an N-window of observations, all of the metrics described above can be computed. Although such analyses can be quite useful, it is also of interest to develop measures which are not dependent on an assumed window size. This desire leads to the consideration of asymptotic measures in which limits as N tends towards infinity are considered.

In [49], for example, asymptotic measures are defined for the case where the difference between the mean sequences is persistently exciting. If this is the case, then the (biased) sample correlation matrix of the difference between the mean sequences exists and is defined by,

$$R_{ij}(k) = \lim_{N\to\infty} (N^{-1}) \sum_{j=0}^{N} [m_i(j+k) - m_j(j+k)]\,[m_i(j) - m_j(j)]^t \qquad (4\text{-}13)$$

Since $m_i(k)$, $C_i(k)$ and $R_{ij}(k)$ are vector and matrix time sequences, we can take Fourier transforms element by element to obtain $m_i(\omega)$, $C_i(\omega)$, and $R_{ij}(\omega)$. The transformed quantities are then used to compute asymptotic distinguishability measures. For example, the asymptotic B distance, is defined by

$$\bar{B}_{ij} = \lim_{N\to\infty} B_{ij}(N) \;/\; N \qquad (4\text{-}14)$$

where $B_{ij}(N)$ is the B distance between hypotheses H_i and H_j based on the N-window of data defined in Eq. 4-10. It is then shown that B_{ij} can be computed using a frequency domain integral. In the one dimensional case with equal covariance functions under H_i and H_j($C_i = C_j = C$) this integral reduces to,

$$\bar{B}_{ij} = 1/4\,(2\pi)^{-1} \int_0^{2\pi} R_{ij}(\omega)\, C^{-1}(\omega)\, d\omega \qquad (4\text{-}15)$$

Some interesting properties of Eq. 4-15 are described in [49]. For example, (Dirac) delta functions in $R_{ij}(\omega)$ can be easily handled. This corresponds to the case where the difference between the means is a sinusoid. The sinusoid is persistently exciting and, therefore, results in a finite limit to Eq. 4-14. Note, however, that since $\bar{B}_{ij}$ is finite, the limit of $B_{ij}(N)$ is infinite.

In the FDI problem we are frequently interested in cases where signals are not persistently exciting and, hence, there is a finite limit to $B_{ij}(N)$. If this is the case, then the asymptotic measure in Eq. 4-14 is zero and provides no information. We now extend the results of [49] for nonpersistent excitation in the one dimensional case with equal covariance functions. The general case is handled in a similar manner.

First expand Eq. 4-15 (without the limits) in terms of the difference between the mean sequences. This gives

$$N^{-1}B_{ij}(N) = K \int_0^{2\pi} d\omega\ C^{-1}(\omega)\{ \sum_{k=0}^{\infty} e^{-j\omega k}\ N^{-1} \sum_{t=0}^{N} [m_i(t+k) - m_j(t+k)][m_i(t) - m_j(t)]\} \tag{4-16}$$

Now multiply both sides of Eq. 4-16 by N^{-1} and change summation variables ($n_1 = t+k$). This gives

$$B_{ij}(N) = K \int_0^{2\pi} d\omega\ C^{-1}\ (\omega) \sum_{n_1=0}^{\infty} \sum_{t=0}^{N} e^{-j\omega n_1}\ e^{j\omega t}[m_i(n_1)-m_j(n_1)]\ [m_i(t) - m_j(t)] \tag{4-17}$$

Taking limits as N goes to infinity we recognize that the double summation is just the product of the Fourier transform of the difference between the means and its complex conjugate. Thus, we have,

$$B_{ij}^{W} = \lim_{N\to\infty} B_{ij}(N) = K \int_0^{2\pi} d\omega\ C^{-1}\ (\omega)\ [m_i(\omega) - m_j(\omega)]^2 \tag{4-18}$$

which is the desired result. This equation holds when the limits exist and this occurs only when the mean sequences have Fourier transforms. The latter requirement implies that both mean sequences be square integrable (i.e., have finite total energy). Thus, the limits do not exist for means which satisfy

the persistent excitation requirement of [49]. If we do admit delta functions, however, then the result is consistent with the original result in that B_{ij}^{w} is infinite. Following [49] and arguments similar to the above for the vector case (again assuming equal covariance functions), Eq. 4-18 becomes,

$$B_{ij}^{w} = \lim_{N\to\infty} B_{ij}(N) = 1/4\ (2\pi)^{-1} \int_{o}^{2\pi} d\omega\ [m_i(\omega)-m_j(\omega)]^t\ C^{-1}(\omega)\ [m_i(\omega) - m_j\ (\omega)] \tag{4-19}$$

EXAMPLE

Let $m_j = 0$, and $m_i(k)$ be a narrow band process with finite energy centered near ω_c (and near $2\pi - \omega_c$ due to sampling) in the direction v_i. For example $m_i(k) = [Ep_N(k)\ \sin(\omega_c k)]v_i$ where $p_N(k)$ is a pulse function of length N (N large). Then we have,

$$B_{ij}^{w} = (E^2/4)\ v_i^t\ C^{-1}(\omega_c)\ v_i \tag{4-20}$$

4.2 EVALUATION METHODS FOR CONSTANT DIRECTION FAILURE SIGNATURES

In this subsection we consider the problem of evaluating the distinguishability of hypotheses of the form,

$$H_i : z(k) = C_i f_i(k) + n(k) \tag{4-21}$$

where z(k) is an n dimensional observation vector, n_k is a stationary Gaussian colored noise process, C_i is an n dimensional failure direction, and $f_i(k)$ is a failure signature. This problem is of interest since many FDI schemes attempt to make decisions based on observations such as Eq. 4-21 (e.g., y_k representing detection filter residuals; the decentralized aircraft-path residuals to be defined in Section 5; or input disturbance estimates in the control element FDI scheme of [50]).

The distinguishability of two hypotheses H_i and H_j may be determined using the metrics in subsection 4.1 when C_i is known, the covariance function of n(k) is given and $f_i(k)$ is adequately characterized (either deterministically or stochastically). However, for the FDI evaluation problem we are more interested in learning what characteristics of $f_i(k)$ are required to achieve some desired level of distinguishability. Thus the methods to be developed below will be seeking information about the "size" (e.g., total energy) and spectral content of the failure signature which is needed to achieve a given level of failure distinguishability.

4.2.1 Detectability

Let the hypothesis H_o, characterized by $C_i = 0$, convey the hypothesis that the system is in normal operation (i.e., no failure). The ability of any FDI system to detect H_i is then computed by examining distinguishability measures between H_o and H_i. Since z(k) is a time series, we can either make use of time domain measures by forming an N-window of measurements, or compute asymptotic measures in the frequency domain. Both cases are considered here.

TIME DOMAIN METHOD 1

In this method we assume that $f_i(k) = Es_i(k)$ and that some particular sequence $s_i(k)$ for k=1, ...N, is of interest. For example, $s_i(k) = 1$ might be used to evaluate how large the average value of $f_i(k)$ needs to get to be detectable; or $s_i(k) = \sin(\omega_c k)$ might be used to evaluate the frequency regions of greatest detectability. Let ζ be defined as in Eq. 4-10 with mean under H_i, m_ζ^i, and covariance matrix under both H_i and H_o, C_ζ. If B_d is a desired level of detectability (determined using the bounds of Eq. 4-6, then we can compute the value of E which achieves B_d from

$$E_{min}^2 = B_d[m_\zeta^t C_\zeta^{-1} m_\zeta]^{-1} \qquad (4\text{-}22)$$

Furthermore it is easy to see that for $E^2 > E_{min}^2$, B is larger implying greater detectability for "larger" failures. Note, however, that many smaller failures may also be frequently detected by an FDI scheme. This is because Eq. 4-22 merely gives the failure size such that both the probability of false alarm and the probability of missed detection are small. Thus, for an FDI test which achieves the desired P_{FA}, smaller failures (than E_{min}) could be detected, but with less reliability (i.e., larger P_{MD}).

TIME DOMAIN METHOD 2

In this method, we also assume that $f_i(k) = Es_i(k)$, but ask the question How large must E be so that the worst case sequence $s_i(k)$, where $s_i(k)$ is constrained to have unit energy, is detectable? Again, using ζ as defined in Eq. 4-10, the covariance matrix under H_i and H_o, C_ζ, is easily computed. The mean vector is expressed by

$$m_\zeta^i = E\, C_N s_N \qquad (4\text{-}23)$$

where C_N is an nN x N matrix defined by,

$$C_N = \begin{bmatrix} C_i & 0 & 0 & \cdots & 0 \\ 0 & C_i & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & & \cdots & & C_i \end{bmatrix} \qquad (4\text{-}24)$$

and s_N is an N x 1 vector defined by,

$$s_N^t = (s_i(1), \ldots, s_i(N)) \qquad (4\text{-}25)$$

The worst case signature with, for example unit energy, is defined as the one which causes B to be a minimum. Thus s_N is found by solving,

$$\min \; 1/8 \; E^2 \; s_N{}^t \; [C_N{}^t C_\zeta{}^{-1} C_N] \; s_N$$

$$\text{s.t. } s_N{}^t s_N = 1 \qquad (4\text{-}26)$$

The solution to Eq. 4-26 is found using Rayleigh's inequality: s_N is the eigenvector of the matrix in brackets corresponding to its smallest eigenvalue, λ_{min}, and the minimum is, therefore, $E^2\lambda_{min}$. Thus, $E_{min}{}^2$, the smallest value of E^2 which gives $B = B_d$ is just,

$$E_{min}{}^2 = 8 \; B_d \; / \; \lambda_{min} \qquad (4\text{-}27)$$

TIME DOMAIN METHOD 3

The last time domain method is directed at evaluating average limits when the failure signature $f_i(k)$ can be described stochastically. As we will see, this method results in extreme conservatism in terms of detectability and is, therefore, not recommended. The extreme conservatism comes from the fact that $f_i(k)$ is described as a zero mean Gaussian process. Although this may in some cases be an accurate description of behavior under the "failure" hypothesis, the utility of labeling undetectable signals (i.e., frequently near zero) as failures is questionable.

In this case, we let $f_i(k) = Es_i(k)$ be described by a first order difference equation,

$$x_f(k) = A_f x_f(k-1) + B_f \; w(k-1)$$

$$s_i(k) = C_f x_f(k) + D_f \; w(k) \qquad (4\text{-}28)$$

where x_f is an m dimensional state vector and w(k) is a Gaussian white noise sequence. Using Eq. 4-28, $f_i(k)$ is a zero mean Gaussian process with autocovariance function $\Sigma_f(k) = E^2\Sigma_s(k)$, where $\Sigma_s(k)$ is the autocovariance function of $s_i(k)$ (computed from Eq. 4-28, e.g., see [22]). Define ζ as in Eq. 4-10, with $m_\zeta = 0$ under H_i and H_o, and with autocovariances,

$$\Sigma_o(k) = E[z(j)z^t(j+k) \mid H_o] \tag{4-29}$$

$$\Sigma_i(k) = E[z(j)z^t(j+k) \mid H_i] = \Sigma_o(k) + C_iC_i^t\Sigma_f(k) \tag{4-30}$$

Plugging these statistics into Eq. 4-8 yields,

$$B = 1/2 \text{ Ln det } \{ [\Sigma_i\Sigma_o^{-1}]^{1/2} + [\Sigma_o\Sigma_i^{-1}]^{1/2}\} \tag{4-31}$$

where Σ_i and Σ_o are the covariance matrices of ζ under H_i and H_o respectively. Equation 4-31 can be simplified using Eq. 4-30 for more efficient calculation. However, its extreme conservatism makes it a poor choice for FDI evaluation. To see this, consider the case where N = 1, $\Sigma_f = 1$ and z(k) is a one dimensional process. For the high signal to noise ratio case ($C_i^2 >> \Sigma_o$), Eq. 4-31 is approximately

$$B \sim 1/2 \text{ Ln } [(C_i^2/\Sigma_o + 1)/2]^{1/2} \tag{4-32}$$

For a reasonable value of B (say 10) Eq. 4-32 implies that C_i must be on the order of 10^9 times greater than $\Sigma_o^{1/2}$ in order to be detectable! Furthermore, the conservatism of Eq. 4-32 is not entirely due to the looseness of the guaranteed performance bounds in Eq. 4-6. In the one dimensional case the value of C_i which achieves any desired P_e (for equal false alarm and missed detection probabilities) can be computed. To achieve $P_e = 10^{-4}$ (corresponding to

about B = 10^{-} for the upper bound in Eq. 4-6) we require that $.01(C_i^2 + \Sigma_o)^{1/2} \sim 2.5\ \Sigma_o^{1/2}$ (this is derived using Eq. 4-1 knowing that the pdfs under the two hypotheses cross at only two points). Thus even using more exact calculations for the one dimensional case requires that C_i be about 250 times as big as $\Sigma_o^{1/2}$ to be detectable. Although this is considerably smaller than Eq. 4-32 would indicate, it is still exceptionally conservative.*

FREQUENCY DOMAIN METHOD 1

Since z(k) is a time series, we can use the asymptotic measures described in subsection 4.1 to evaluate detectability. In none of these methods will $f_i(k)$ be described as a zero mean process due to the discussion above. However, such methods are easily derived. In the first method, as in the first time domain case, we assume that some aspect of $f_i(k)$ is known. To use the frequency domain integral in Eq. 4-19, however, it is evident that only the squared Fourier transform of $f_i(k)$ (its power spectrum) needs to be specified. Thus, this analysis is not specific to a particular signature, but is valid for the class of failure signatures which have the same power spectrum. Let $S_n(\omega)$ denote the spectral density of n(k), and $f_i(k) = Es_i(k)$ with $s_i(k)$ a unit energy signal with power spectrum $m_s^2(\omega)$ (i.e., $\int_0^{2\pi} m_s^2(\omega)\, d\omega = 1$). Then, to achieve $B^W = B_d$, Eq. 4-19 requires,

$$E_{min}^2 = B_d / \left[1/4(2\pi)^{-1} \int_0^{2\pi} m_s^2(\omega)\ C_i^t S_n^{-1}(\omega)\ C_i\ d\omega\right] \qquad (4\text{-}33)$$

*Note that in the one dimensional nonzero-mean case, calculations using B distance instead of actual error probabilities result in estimates of signature size which are about a factor of 9 too big.

Candidates for $m_s(\omega)$ include wideband processes (e.g., the Fourier transform of e^{-ak}, or $p_N(k)$; a pulse of length N) and narrowband processes (e.g., the Fourier transform of $p_N(k)\sin(\omega_c k)$). For wideband processes it may be of interest to plot the integrand of the denominator of Eq. 4-33 vs. ω to highlight the regions in the frequency domain which contribute most to detectability. For the narrow band processes, plots of E_{min} vs. ω_c also indicate frequency regions which are highly detectable.

FREQUENCY DOMAIN METHOD 2

Paralleling the development of time domain methods above, we again assume $f_i(k) = Es_i(k)$ and find E such that the worst case signature spectrum, $m_s(\omega)$, is detectable. The worst case signature minimizes B^W subject to the unit energy constraint on $s_i(k)$. Thus we must first solve,

$$\min \int_0^{2\pi} E^2 m_s^2(\omega)\; C_i{}^t S_n^{-1}(\omega)\; C_i \; d\omega \qquad (4\text{-}34a)$$

$$\text{s.t.} \int_0^{2\pi} m_s^2(\omega)\; d\omega = 1 \qquad (4\text{-}34b)$$

Since all factors in the integrand of Eq. 4-34a are positive, the solution is for $m_s(\omega)$ to be a narrow band process at the frequency for which $C_i{}^t S_n^{-1}(\omega)\; C_i$ is minimum. The minimally detectable energy is then computed as in Eq. 4-33. This solution is, unfortunately, of little use in practical problems since S_n must be estimated and, therefore, the exact frequency which results in minimal detectability is of little interest. The dual problem to Eq. 4-34, however, may be of interest. That is,

$$\min E^2 \int_0^{2\pi} m_s{}^2(\omega)\, d\omega$$

$$\text{s.t. } 1/4\ (2\pi)^{-1} \int_0^{2\pi} E^2 m_s{}^2(\omega)\ C_i{}^t S_n{}^{-1}(\omega)\ C_i\ d\omega = B_d{}^W \qquad (4\text{-}35)$$

The solution to Eq. 4-35 is not immediately obvious. However, a similar problem can be stated which has a simple solution. Let $f(\omega) = E^2\ m_s{}^2(\omega)$, and $g(\omega) = C_i{}^t S_n{}^{-1}(\omega)\ C_i$. The constraint in Eq. 4-35 can now be rewritten using inner product notation to denote the integral over the interval [0, 2π], as

$$\langle f,g \rangle = B_d{}^W \qquad (4\text{-}36)$$

The minimum norm function, f, which achieves Eq. 4-36 can be easily found. It is,

$$f = [B_d{}^W / \langle g,g \rangle\ ]\ g \qquad (4\text{-}37)$$

Thus, we have,

$$f^*(\omega) = [C_i{}^t S_n{}^{-1}(\omega)\ C_i] B_d{}^W / \left[1/4\ (2\pi)^{-1} \int_0^{2\pi} (C_i{}^t S_n{}^{-1}(\omega)\ C_i)^2 d\omega\right] \qquad (4\text{-}38)$$

as the solution to

$$\min \int_0^{2\pi} f^2(\omega)\, d\omega$$

$$\text{s.t. } 1/4\ (2\pi)^{-1} \int_0^{2\pi} f(\omega)\ C_i{}^t S_n{}^{-1}(\omega)\ C_i\ d\omega = B_d{}^W \qquad (4\text{-}39)$$

The energy in the signature f(k) needed to achieve Eq. 4-37 is computed from Eq. 4-36 using

$$E_{min}{}^2 = \int_0^{2\pi} f^*(\omega)\, d\omega \qquad (4\text{-}40)$$

4.2.2 Distinguishability

Failure distinguishability refers to the inherent ability to distinguish failure modes from each other (not including the no-failure mode). Although any of the calculations discussed in subsection 4.1 provides, in principle, a direct means for evaluating failure distinguishability , there is a problem in that the signature energies for different failure modes must be specified. One logical approach might be to use the minimally detectable energies determined from a detectability analysis such as those described above. However, it is possible to show that such an approach does not guarantee that larger failure signature energies are more distinguishable. In order to provide a more meaningful distinguishability analysis in this regard, we define a new distinguishability concept.

As is well known in the statistical literature, the assumption of a particular statistical behavior for an observation under some hypothesis, $\bar{H}_i$, allows significance only in the testing of the null hypothesis, H_i. That is, given a model for H_i, we can only design a test which reliably rejects that hypothesis. Now suppose we wish to define a region in measurement space which represents a highly significant decision that the i^{th} control element is not failed. Furthermore, suppose we wish to define this decision region so that the resulting test is invariant with respect to the failure signature energy E_i.

As in subsection 3.3, let the α-significance level test for a known signal energy (consider one sample of z(k)) be defined by

$$D(E_i) = \{Z : Pr [\|z(k) - C_i f_i(k)\|^2 > \|Z - C_i f_i(k)\|^2] < \alpha\} \qquad (4\text{-}41)$$

The E_i invariant decision region is then the locus of points in measurement space defined by Eq. 4-40 for all possible E_i, or

$$D = \bigcap_{E_i} D(E_i) \tag{4-42}$$

Thus, as in subsection 3.3, the decision region, D, is defined by projecting the α-significance contour for any E_i into the null space of C_i.

We can now define the minimum signature energy needed to distinguish failure i from failure j, $E_{i/j}$, as the minimum signature energy required to achieve some desired value of a given metric (e.g., B^W) after projection into the null space of failure j. Notice that this definition is not symmetric (i.e., $E_{i/j} \neq E_{j/i}$). Furthermore, note that this analysis quantifies the limits to performance of the "isolation" hypothesis tests discussed in subsection 3.3

SECTION 5

CONTROL ELEMENT FDI DESIGN AND EVALUATION FOR THE B-737 AIRCRAFT

In this section, the ideas described in Sections 3 and 4 are applied to the control element failure detection and isolation problem. Aircraft-path and actuator-path failures are considered in independent subsystems as discussed in Section 2. For both subsystems we have assumed that all sensors have been validated (i.e., contain only "in-spec" errors). For the actuator-path subsystem, multiple simultaneous and sequential failures are allowed. For the aircraft-path subsystem, only single failures are allowed. Also, no aerodynamic effects other than those due to control element failures are allowed (note: such effects are likely to be detected, however, isolation performance is severely affected).

Both subsystems make use of the "trigger/verify/(isolate)" structure discussed in subsection 3.1 for handling the unknown onset-time problem. In the aircraft-path subsystem, we have concentrated heavily on developing a formal design methodology which allows the designer to assess FDI performance capabilities and limitations as a function of both sensor noise and model uncertainty. This was done to convert, as much as possible, the typical algorithm tuning process into one of validating model error assumptions. Except for defining threshold scheduling procedures, this has been successfully

accomplished. For the actuator path subsystem, two separate hypothesis testing algorithms were examined. The first algorithm was designed using only static thresholds which are selected on the basis of average error budgets. The second algorithm used threshold schedules which assumed that all error sources could be bounded by a high-pass transfer function error.

In the remainder of this section, we will detail the design, development, and evaluation of both aircraft-path and actuator-path control element FDI systems. Evaluation of the algorithms was accomplished using data generated from NASA's modified B-737 simulation. This simulation is a six degree of freedom simulation with nonlinear aerodynamic models, realistic actuation models (high order dynamics, rate limits, cable stretch, etc.), a Dryden wind model, and sensor errors including white noise, biases, and scale factors. In the case of aircraft-path failures, we will also evaluate fundamental limits to FDI performance based on flight test data obtained from NASA's Transportation Systems Research Vehicle (TSRV).

5.1 AIRCRAFT PATH FAILURE DETECTION AND ISOLATION

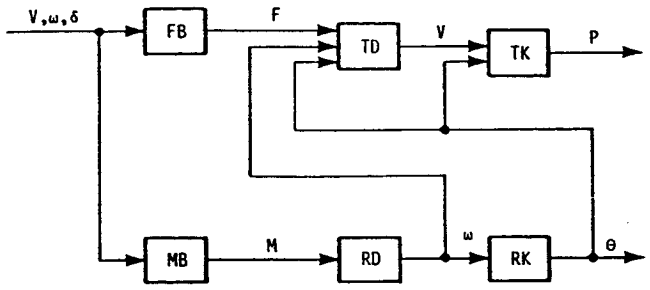
The aircraft path FDI problem described in Section 2 is an attempt to characterize a very broad class of failure modes for which a single FDI system is desired. The failure modes considered as aircraft path failures are those failures in which the effective control value (i.e., the value of unfailed control "deflection" which results in the same applied aerodynamic forces and moments as the failure mode) is different than the measured control value. Thus, any failure occurring outboard of the measured control value is considered as an aircraft path failure. This includes the traditional "partially missing" aircraft path failure mode as well as stuck, floating, runaway, etc.,

when these failures occur in the aircraft path. To achieve such generality in the FDI system, specific temporal characteristics of failure signatures will not be used explicitly. Thus, systems such as [51] are not of interest. The OSGLR method described in [10] provides insensitivity to failure signature information through estimation of temporal basis coefficients. In [52] this method was applied using standard Kalman filtering techniques to generate residuals. However, the method is also applicable to other types of residuals. In our approach, we make a broad assumption that the failure signature is coherent. That is, we assume that the signatures of important failures remain large and of one sign throughout the length of time needed to detect and isolate failures. This assumption leads to specific structures for the FDI process. We then concentrated our efforts in developing analytical design and evaluation methodologies for selecting parameters and predicting performance. In this section we detail the application of these methodologies.

5.1.1 Decentralized Residual Generation

The first step in designing a system to generate residuals for aircraft path FDI is the assessment of the various sources of analytic redundancy which are available.

Figure 5-1 shows the interrelationships of the various elements which describe the dynamics of a rigid body aircraft. The vector equations and variable definitions are shown in Table 5-1 and are derived in [29]. All quantities are in c.g. -- centered body-axis coordinates (unless specified otherwise). (In addition to these elements one might include relationshps between actuation effort (hinge moment) and V, δ, ω, and g as a potential source of redundancy. However, these are typically unreliable [53].



KEY:

FB = FORCE BALANCE
MB = MOMENT BALANCE
TD = TRANSLATIONAL DYNAMICS
RD = ROTATIONAL DYNAMICS
TK = TRANSLATIONAL KINEMATICS
RK = ROTATIONAL KINEMATICS

Figure 5-1. Generalized Rigid Body Aircraft Model

TABLE 5-1. ELEMENTS OF AN AIRCRAFT MODEL

FB: $F_a = f(V_w, \omega_w, \delta)$

MB: $M_a = g(V_w, \omega_w, \delta)$

TD: $m(\dot{V} + \omega \times V) = F_a + m\underline{g}$

RD: $I\dot{\omega} + \omega \times I\omega = M_a$

TK: $\dot{P} = R_B^I(\theta)\, V$

RK: $\dot{\theta} = H(\theta)^{-1}\, \omega$

F_a = Aerodynamic force vector
M_a = Aerodynamic moment vector
V = Translational velocity vector
ω = Angular velocity vector
P = Position vector in ECEF coordinates
θ = Euler Angles
V_w = Translational velocity of aircraft relative to air mass $\cong V$
ω_w = Rotational velocity of aircraft relative to air mass = ω
δ = Surface deflections
$\underline{g}$ = Acceleration of gravity
m = Mass
I = Moment of inertia matrix
R_B^I = Rotation matrix from body-axes to ECEF axes
H = Transformation matrix (see [10])

Although several aircraft relationships could be considered in the residual generation process, it is the force and moment balances which provide the most useful information about aircraft path failures. This is because these relations are the only ones affected by failures and because other relations have potential for adding only uncertainty due to model error (e.g., effects of wind acceleration on TD).

If measurements of both inputs and outputs of these relations were available, static or memoryless residuals could be generated. For the force balance relation, this is the case because accelerometer measurements are available. That is, let A_m be the measured specific force from three orthogonal accelerometers centered along the body axes. Then we have

$$A_m = \dot{V} + \omega \times V - g \tag{5-1}$$

The force balance equation in Table 5-1 then allows the formation of an independent estimate of these measurements from air data, angular velocity, and deflection measurements. A set of three translationa residuals can then be defined by,

$$\nu_T = \begin{matrix} \nu_X \\ \nu_Y \\ \nu_Z \end{matrix} = A_m - f(V), \omega, \delta)/m \tag{5-2}$$

For the moment balance relation, memoryless residual generation would require direct measurement of aerodynamic moments. Since these are not available we will use the rotational dynamic relations in addition to the moment balance equations to form finite memory "rotational" residuals as follows. First, write a discrete time, nonlinear, time invariant state space model for the aircraft, viz.,

$$x(k+1) = a[x(k), \delta(k)] \tag{5-3}$$

where x(k) is the n-dimensional state vector and $\delta(k)$ is the m-dimensional input vector. The moment balance and rotational dynamic relationships give rise to the three components of Eq. 5-3 corresponding to the angular velocity states, $\omega(k)$. Since we have complete measurements of the states, x, and inputs, δ, for these three equations, three rotational residuals can be defined by,

$$\nu_R = \begin{pmatrix} \nu_P \\ \nu_Q \\ \nu_R \end{pmatrix} = \omega(k+1) - a_\omega[x(k), \delta(k)] \tag{5-4}$$

where $a_\omega(\cdot)$ is the component of $a(\cdot)$ which corresponds to the states, ω.

The rotational and translational residuals defined above are clearly decentralized in that they only use those parts of the system model which contain direct information about the failures of interest. In an ideal world, the other relations provide the information which would be necessary to optimally incorporate all measurements into the residual generation process through use of a Kalman filter. However, errors in these relationships can make such an approach suboptimal for the failure detection problem. With the decentralized approach we are sure that the errors in unused relations will not have any effect on FDI performance and for those errors which do, analytical evaluation of their impact will be possible.

The use of static and finite memory residual formulations is chosen for the simplicity of failure characterization. In particular, we see that, for a broad class of force and moment relationships, failures will show up in fixed directions in the residual space defined by Eqs. 5-2 and 5-4. For this to be the case, the functions $f(\cdot)$ and $a_\omega(\cdot)$ must be of the form,

$$f(\cdot) = f_1(x) + \sum_i b_i \, g_i(\delta_i) \tag{5-5}$$

$$a_\omega(\cdot) = a_1(x) + \sum d_i g_i(\delta_i) \tag{5-6}$$

where, again, x is a vector of measureable states needed in Eqs. 5-2 and 5-4. The failure direction corresponding to the i^{th} control element failure in the residual space defined by $\nu^t = (\nu_T{}^t, \nu_R{}^t)$ is then $C_i{}^t = (b_i{}^t, d_i{}^t)$. Although Eqs. 5-5 and 5-6 may seem restrictive, they are in fact typical of standard aircraft models (i.e., linear aerodynamic models). The details of these equations for the B-737 are now given.

The standard non-dimensional description of forces and moments referenced to a body fixed coordinate frame takes the form

$$X = \overline{Q} \cdot S \cdot CXB \tag{5-7}$$

$$Y = \overline{Q} \cdot S \cdot CYB \tag{5-8}$$

$$Z = \overline{Q} \cdot S \cdot CZB \tag{5-9}$$

$$L = \overline{Q} \cdot S \cdot b \cdot CLB \tag{5-10}$$

$$M = \overline{Q} \cdot S \cdot \overline{c} \cdot CMB \tag{5-11}$$

$$N = \overline{Q} \cdot S \cdot b \cdot CNB \tag{5-12}$$

where $\overline{Q}$ represents the dynamic pressure (in units of lbs/sq. ft.), S represents a reference surface area (sq. ft.), b is the reference wing span (ft.), and c is the reference mean chord length (ft.). The non-dimensional coefficients (CXB, CYB, CZB, CLB, CMB, CNB) at any point in time are typically related to the instantaneous values of control surface deflections, engine throttle setting, relative velocity and direction of the air mass, and inertial angular

rates. In some cases, air mass acceleration is included using angle of attack rate, α. For our purposes we will assume that these relationships are primarily linear (except in the case of the rotary terms p, q, r where total velocity is used as a divisor to obtain non-dimensional coefficients).

Thus, neglecting terms which are typically close to zero, the six nondimensional coefficients are assumed to be in the following form.

$$CXB = \frac{\bar{c}}{2V} C_{Xq} \cdot q + C_{X\alpha} \cdot \alpha + K_X + \sum_i C_{X\delta_i} \delta i \qquad (5\text{-}13)$$

$$CYB = C_{Y\beta} \beta + \frac{b}{2V} [C_{Yp} p + C_{Yr} r] + K_Y + \sum_i C_{Y\delta_i} \delta i \qquad (5\text{-}14)$$

$$CZB = C_{Z\alpha} \alpha + \frac{\bar{c}}{2V} [C_{Zq} q] + K_Z + \sum_i C_{Z\delta_i} \delta i \qquad (5\text{-}15)$$

$$CLB = C_{L\beta} \beta + \frac{b}{2V} [C_{Lp} p + C_{Lr} r] + K_L + \sum_i C_{L\delta_i} \delta i \qquad (5\text{-}16)$$

$$CMB = C_{M\alpha} \alpha + \frac{\bar{c}}{2V} [C_{Mq} q] + K_M + \sum_i C_{M\delta_i} \delta i \qquad (5\text{-}17)$$

$$CNB = C_{N\beta} \beta + \frac{b}{2V} [C_{Nr} r + C_{Np} p] + K_N + \sum_i C_{N\delta_i} \delta i \qquad (5\text{-}18)$$

where,

V = total velocity (f/s)

α = angle of attack (radians)

β = sideslip angle (radians)

p = roll rate (r/s)

q = pitch rate (r/s)

r = yaw rate (r/s)

and where δ_i represents the 11 control elements including left and right ailerons, stabilizers, elevators, spoilers, throttles, and a rudder. All control element values are measured in degrees except throttles which are measured in kilopounds of thrust. Appendix A provides a derivative of the coefficients in Eqs. 5-13 to 5-18 from linear models.

Thus, the three translational residuals are

$$\nu_x = A_x^m - \hat{X}/m \tag{5-19}$$

$$\nu_y = A_y^m - \hat{Y}/m \tag{5-20}$$

$$\nu_z = A_z^m - \hat{Z}/m \tag{5-21}$$

where A_x^m, A_y^m, A_z^m are the c.g.-centered body-referenced accelerometer measurements, and $\hat{X}$, $\hat{Y}$, $\hat{Z}$ are estimates of the aerodynamic forces (also c.g.-centered body-referenced) obtained by using measured values of the states and control elements in Eqs. 5-7 - 5-9, and 5-13 - 5-15. The aircraft mass is denoted by M.

The rotational dynamic relations (neglecting cross products of inertia) are:

$$I_x \dot{P} + (I_z - I_y)\, RQ = L \tag{5-22}$$

$$I_y \dot{Q} + (I_x - I_z)\, PR = M \tag{5-23}$$

$$I_z \dot{R} + (I_y - I_x)\, QP = N \tag{5-24}$$

where I_x, I_y, I_z = moments of inertia about the body axes (slug - ft^2).

Three estimates of angular acceleration are then

$$\hat{\dot{p}} = \frac{I_y - I_z}{I_x} rq + \frac{\hat{L}}{I_x} \tag{5-25}$$

$$\hat{\dot{q}} = \frac{I_z - I_x}{I_y} rq + \frac{\hat{M}}{I_y} \tag{5-26}$$

$$\hat{\dot{r}} = \frac{I_x - I_y}{I_z} qp + \frac{\hat{N}}{I_z} \tag{5-27}$$

where L, M, N are obtained from using measured values in Eqs. 5-10 - 5-12 and 5-16 - 5-18.

Using the trapezoidal rule for approximating the integrals of Eqs. 5-25 - 5-27 over a single time interval of Δ results in the residual equations,

$$\nu_p = p(k) - [p(k-1) + \frac{\Delta}{2} (\hat{\dot{p}}(k) + \hat{\dot{p}}(k-1)] \tag{5-28}$$

and similarly for q and r. Note that if Eq. 5-28 was scaled by $1/\Delta$, these residuals will have units of angular acceleration (r/s^2).

Finally, for completeness, we note that preprocessing of the accelerometer data and α and β vane measurements may be necessary to obtain the c.g. centered, body-referenced measurements required in the above. Letting (ℓ_x, ℓ_y, ℓ_z) denote the coordinates of any sensor in the desired c.g. centered body axis coordinate frame, then accelerometer compensation takes the form,

$$A_x^{Comp} = A_x^m - [-\ell_x(r^2 + q^2) + \ell_y(pq - \dot{r}) + \ell_z(pr + \dot{q})] \tag{5-29}$$

$$A_y^{Comp} = A_y^m - [\ell_x(pq + \dot{r}) - \ell_y(p^2 + r^2) + \ell_z(rq - \dot{p})] \tag{5-30}$$

$$A_z^{Comp} = A_z^m - [\ell_x(rp - \dot{q}) + \ell_y(rq + \dot{p}) - \ell_z(q^2 + p^2)] \tag{5-31}$$

Compensation of the α and β vane measurements takes the form,

$$\alpha^{Comp} = \alpha^{m} + \ell_x(q/V) - \ell_y(p/V) \qquad (5\text{-}32)$$

$$\beta^{Comp} = \beta^{m} - \ell_x(r/V) - \ell_z(p/V) \qquad (5\text{-}33)$$

Derivation of these equations appears in [29].

Equations 5-19 through 5-21 and 5-28 and the associated roational dynamics and force and moment balances provide the information necessary to generate the six desired residuals. Unfortunately, for this project, the nondimensional coefficients of the B-737 aircraft were not available. Linear models were available, however. Recognizing that there is a one-to-one correspondence between the nondimensional coefficients in Eqs. 5-13 - 5-18 and any linear model, it is possible to derive the nondimensional coefficients from these linear models. This derivation is given in Appendix A along with a comparison of actual nondimensional coefficient values for two flight conditions.

5.1.2 Detectability of Aircraft Path Failures Using Flight Test Data

The availability of flight test data from NASA's B-737 aircraft provided a unique opportunity for realistic evaluation of the feasibility of performing control element FDI using aerodynamic models of the aircraft. This is because many errors which are difficult to simulate are excited during actual flight (e.g. wing bending, nonstationary inputs) and realistic values of errors which are normally accounted for (e.g. sensor noise) are present. Unfortunately, sufficient time was not available to perform a complete analysis. Such analysis would require a flight test program to guarantee that all important error sources were excited over a full range of operating conditions. Thus, this section is meant to demonstrate the application of the concepts developed

in Section 4. The results which follow should be considered optimistic due to the fact that only a limited amount of data was processed. In particular, we used only the data from the first 50 seconds of TSRV flight R380, Run 12JR. During this interval the aircraft performed a mild turn maneuver at approximately constant altitude and throttle setting. Mild turbulence is believed to have been present. Figures 5-2a to 5-2q show traces of interesting quantities for this maneuver. The flight condition used to obtain linear models corresponding to this data is defined by:

V = 160 KIAS

h = 3500 ft

Gear up

Flaps = 15°

γ = 0

The evaluation of detectability presented here is aimed at evaluation of limits to performance using the decentralized residuals described in 5.1.1. Although the general framework presented in Section 4 would, in principle, allow an assessment which is independent of residual generation mechanization, such an assessment would be much more difficult due to the complex nature of the effects of failures on the measured quantities. In contrast, the effects of failures on decentralized residual is easy to characterize resulting in a greatly simplified analysis. Furthermore, decentralized residual generation was required for further development and this analysis proved useful in the FDI design procedure.

The basic premise behind the results of this section is that the residuals could be characterized by the alternate failure (and no failure) hypotheses

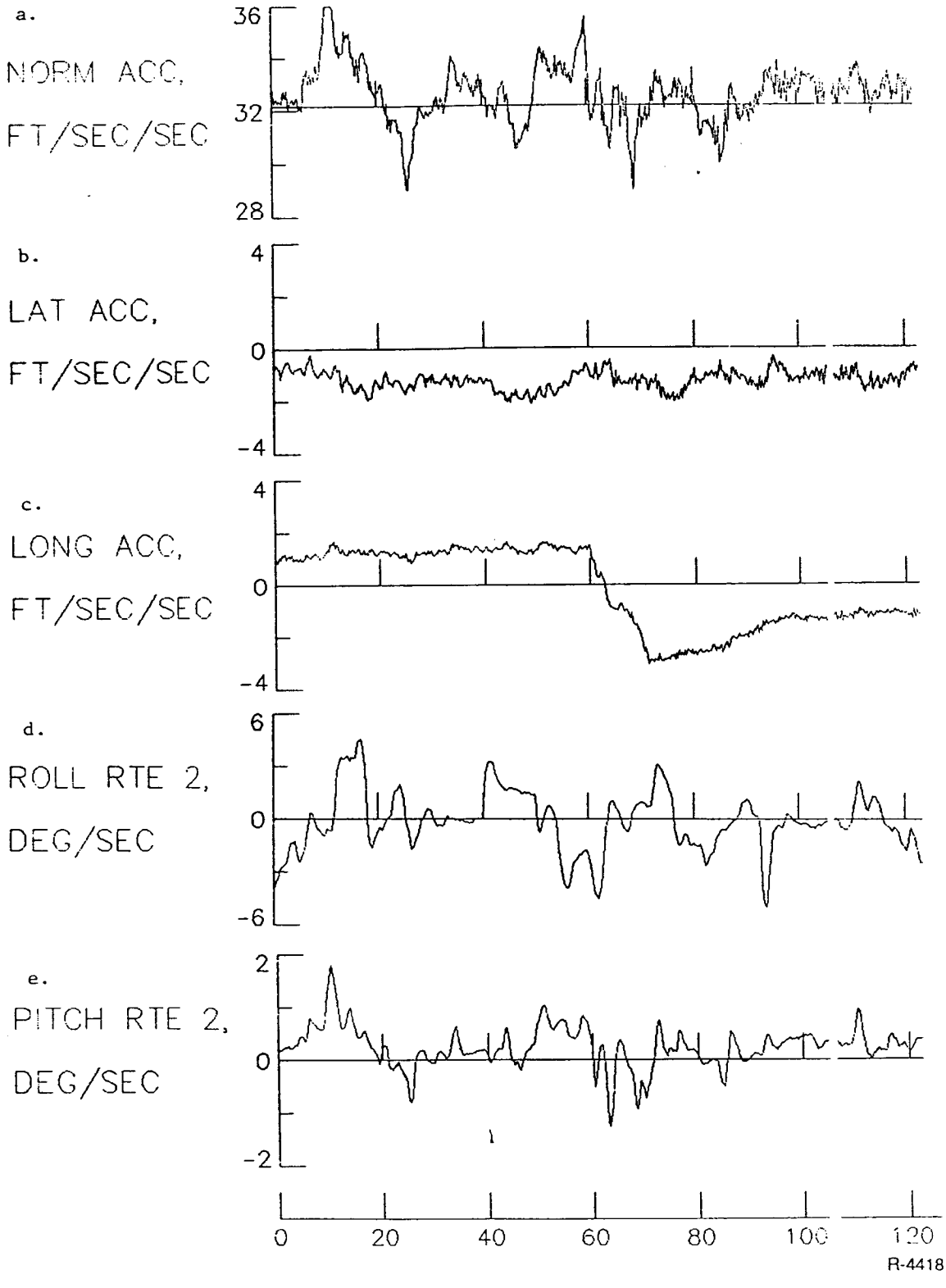


Figure 5-2 (a-q). Flight Data From TYRV R380, Run 12JR

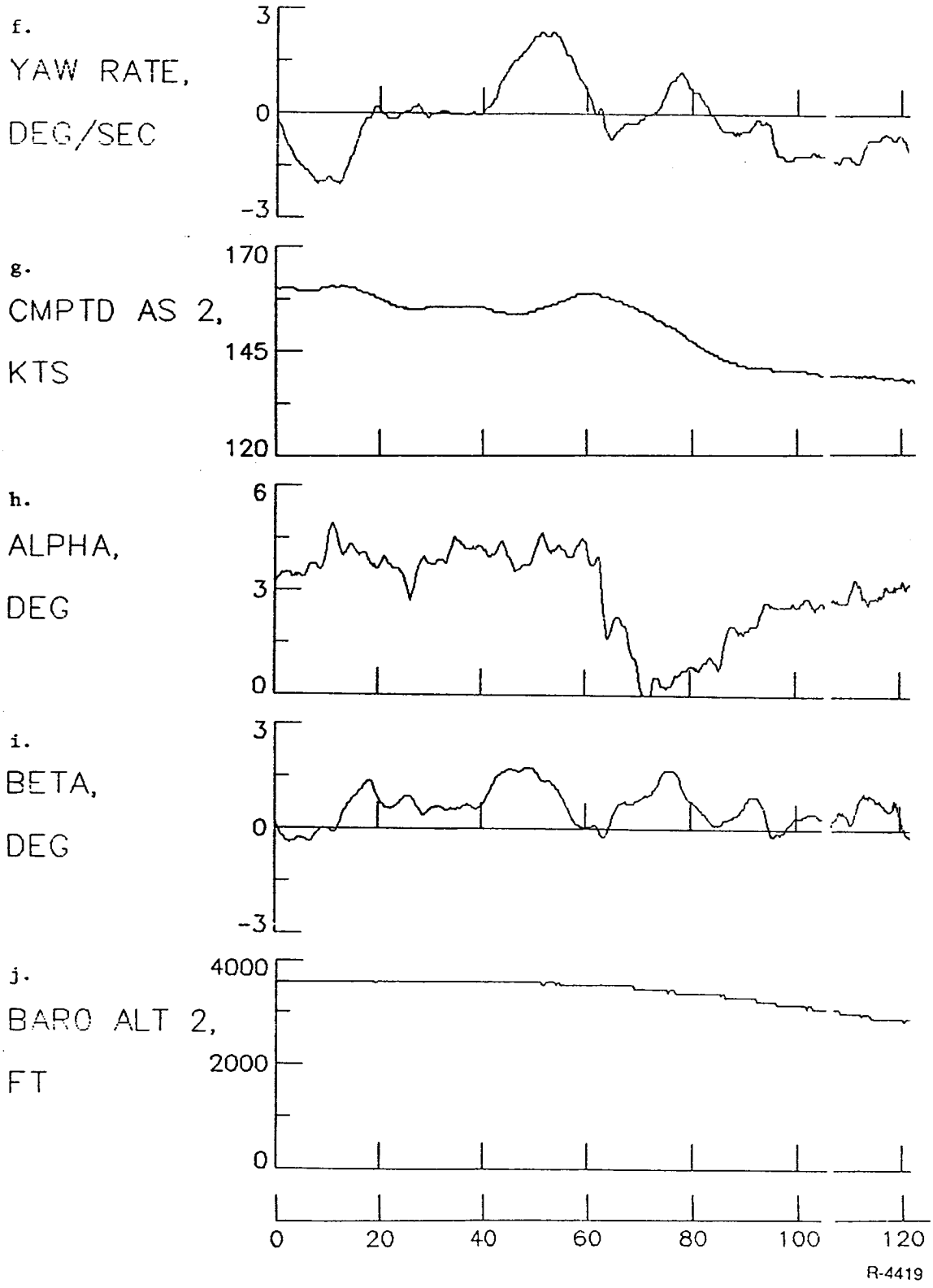


Figure 5-2 (a-q) (Continued). Flight Data From TYRV R380, Run 12JR

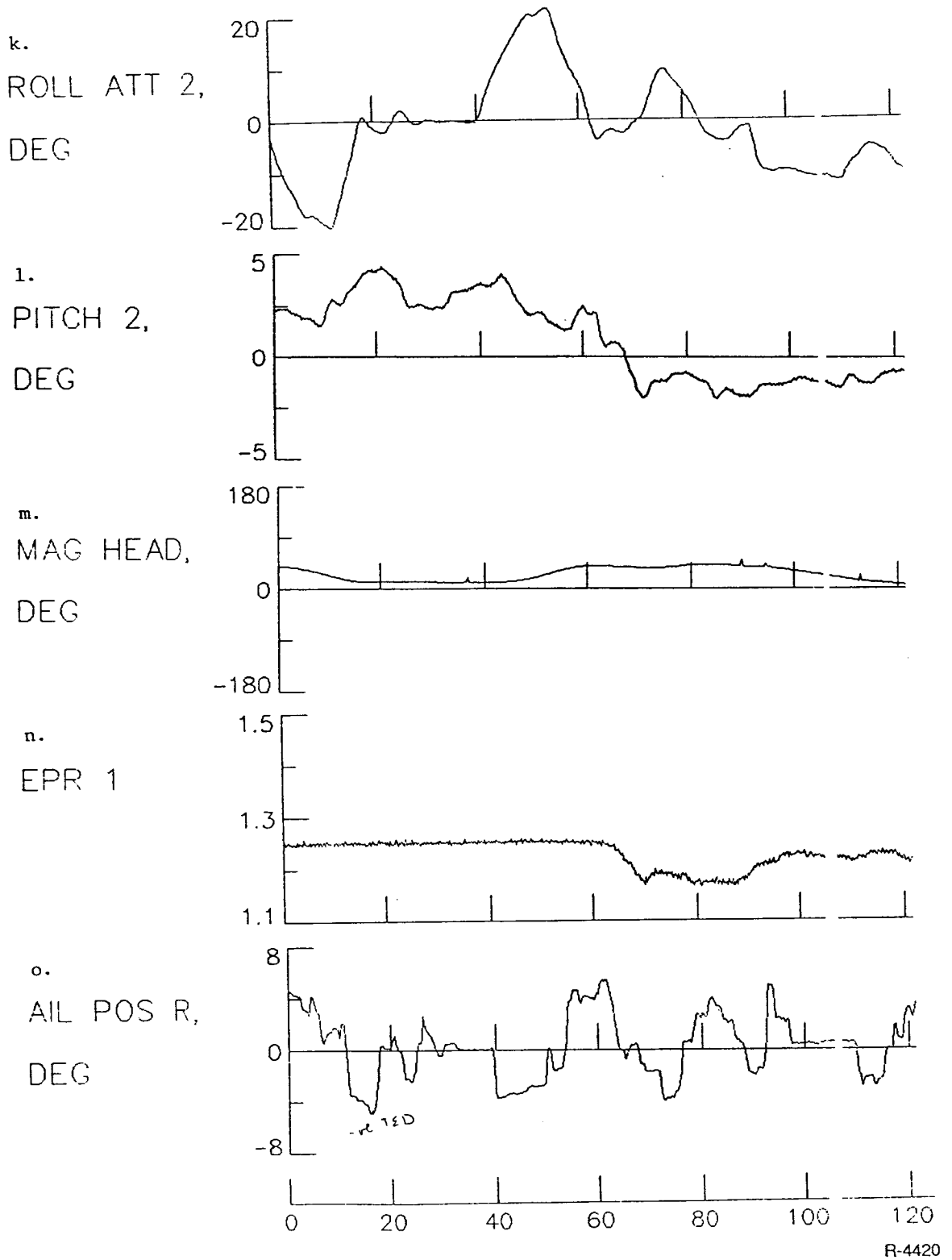


Figure 5-2 (a-q) (Continued). Flight Data From TYRV R380, Run 12JR

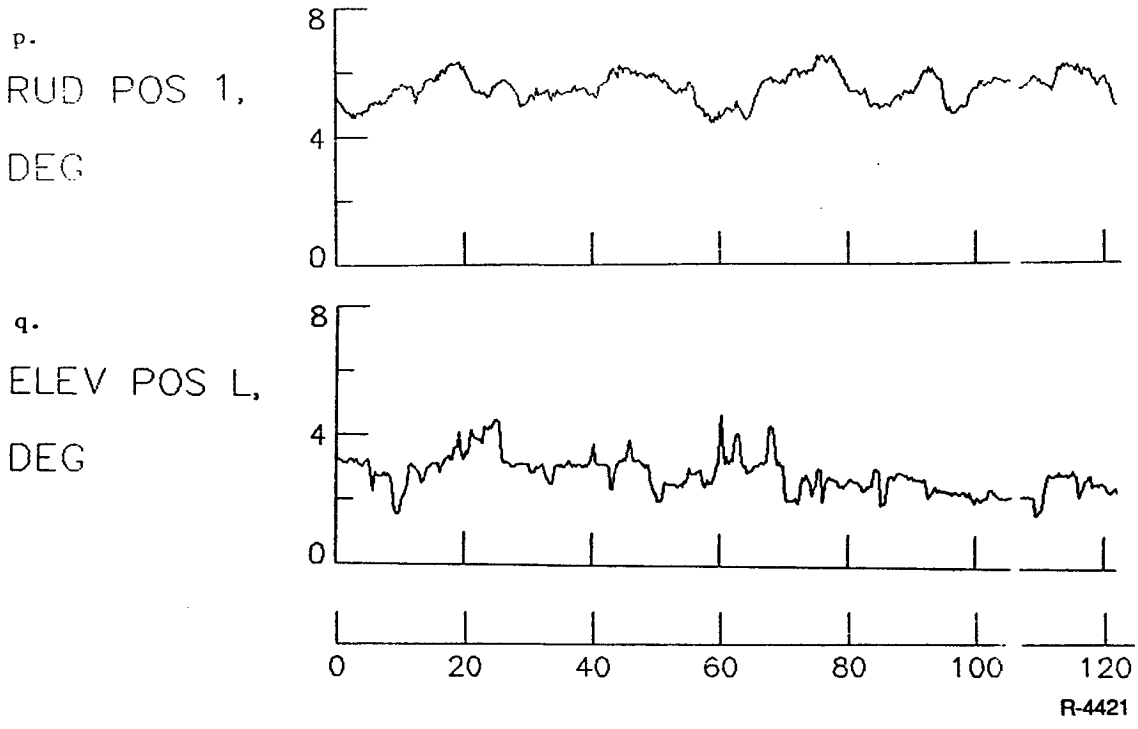


Figure 5-2 (a-q) (Continued). Flight Data From TYRV R380, Run 12JR

described in Equation 4-21. That is, failures show up in fixed directions, modulated by a failure signature, and the errors could be described by a colored noise process whose statistical behavior is independent of the system status (i.e., the same for all hypotheses). This latter property is critical in enabling us to infer the statistical behavior of the residuals under failed modes from the data obtained from normal flight. To see that this is a reasonable assumption consider a model for the decentralized residual of the form,

$$\nu(k) = \eta_0(k) + \sum_{i=1}^{p} (C_i - \hat{C}_i)\delta_i^E(k) + \hat{C}_i[\delta_i^E(k) - \delta_i^m(k)] \qquad (5\text{-}34)$$

where η_0 is a random noise term and where, for each of the p control elements, δ_i^m is the measured deflection, δ_i^E is the effective deflection, $\hat{C}_i$ is the vector of dimensional control derivatives used in generating residuals, and C_i is the actual dimensional control derivative. Under the j^{th} failure hypothesis (j=0 indicates no failure) we have

$$\delta_i^E = \delta_i^m \quad \text{for } i \neq j \qquad (5\text{-}35)$$

$$\delta_i^E \neq \delta_i^m \quad \text{for } i = j \qquad (5\text{-}36)$$

The random noise term, η_0, takes into account the effects of sensor noise as well as modeling error in the formation of the residuals and is, therefore, in general, a nonwhite (and generally nonstationary) process which must be characterized. The term $(\delta_i^E - \delta_i^m)$ in Eq. 5-34 is termed the failure signature and is denoted by $f_i(k)$. The failure signature is important because when it is large, failures are observable. Furthermore, when $f_i(k)$ is large, the

undesirable effect on the aircraft is most pronounced so that we must have a control system which is tolerant of the resulting disturbance vector $\hat{C}_i f_i(k)$ when $f_i(k)$ is small. The results of this section quantify "small" in terms of failure detectability.

Let us now define the behavior of the residuals under no failure by a stochastic process, $\eta(k)$, viz.,

$$H_o : \nu(k) = \eta(k) = \sum_{i=1}^{p} (C_i - \hat{C}_i)\delta_i^m(k) + \eta_0(k) \qquad (5\text{-}37)$$

A simplifying approximation to Eq. 5-34 for the i^{th} failure hypothesis ($i \neq 0$), is then,

$$\nu(k) = \eta(k) + \hat{C}_i f_i(k) \qquad (5\text{-}38)$$

Equation 5-38 assumes that the statistical behavior of δ_i^E, under the i^{th} failure hypothesis, is the same as δ_i^m and that this statistical behavior is independent of the failure mode. Clearly, there are many specific failure mechanisms in which this is not the case. The assumption is nevertheless, very useful in our general failure mode context. For any specific failure mechanism, one must decide whether the results derived here are conservative or optimistic due to this approximation. The advantage of Eq. 5-38 is that the process $\eta(k)$ is observable from flight recorded data (during normal flight) and can, therefore, be realistically characterized.

Having justified the model of Eq. 4-21 for the decentralized residuals we now evaluate the detectability of control element failures using the asymptotic measure B^W defined in Section 4. The procedure is as follows.

1. Compute the six residual signals from Eqs. 5-7 through 5-33 using data recorded during flights of the NASA TRSV.

2. Compute estimates of the power spectral density (PSD) matrix for the six dimensional residual process.

3. Determine the (dimensionalized) failure directions from the aircraft model.

4. Determine the smallest signature energies, for each control element failure, which are required to make specific signature spectra detectable. Plot signature energy versus frequency for narrow band signatures and determine total signature energy for a broadband signature.

RESIDUAL GENERATION

Figures 5-3 through 5-8 show the decentralized residuals computed using TSRV data sampled at 20 Hz (the figures show a lower sampling rate). Sensor compensation (see 5.1.1) was used to accommodate lever arm effects and other sensor validation, averaging, and scaling procedures were used (see Appendix C. Next to each residual is its sample autocovariance function (ACF). The sample ACF is computed from,

$$R(m) = 1/N \sum_{n=1}^{N-|m|} [\nu(n) - \bar{\nu}] [\nu(n+m) - \bar{\nu}]^t \qquad (5\text{-}39)$$

$$\bar{\nu} = 1/N \sum_{n=1}^{N} \nu(n) \qquad (5\text{-}40)$$

for lags m=0, 1, ... 100. For negative lags we use the property $R(m) = R^t(-m)$. Note that R(m) is a biased estimate of the underlying ACF but typically has smaller mean square error (see [59]). Also, for the 50-second data set sampled at 20 Hz, N=1000. The ACFs shown in Figs. 5-3 to 5-8 are the diagonal elements of R(m). There are several salient features of the residuals which are evident from the figures. First, all of the residuals have significantly nonzero means. This is expected due to the inaccuracy in predicting the basic aerodynamic forces and moments of the airframe and due to the nonzero average values of many of the measurements coupled with inaccuracy associated with the

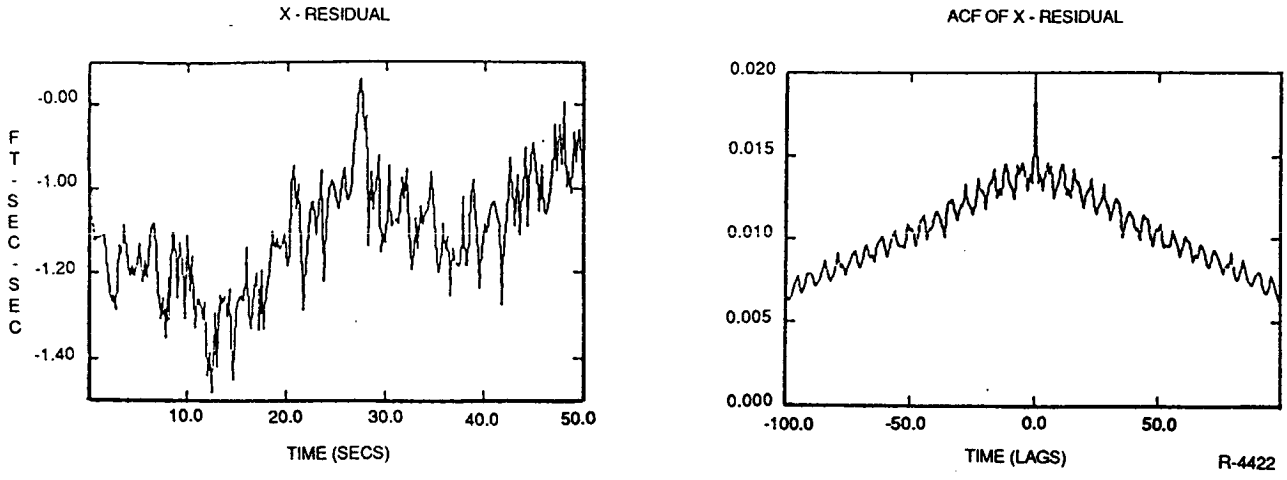


Figure 5-3. Translational X-Residual and ACF

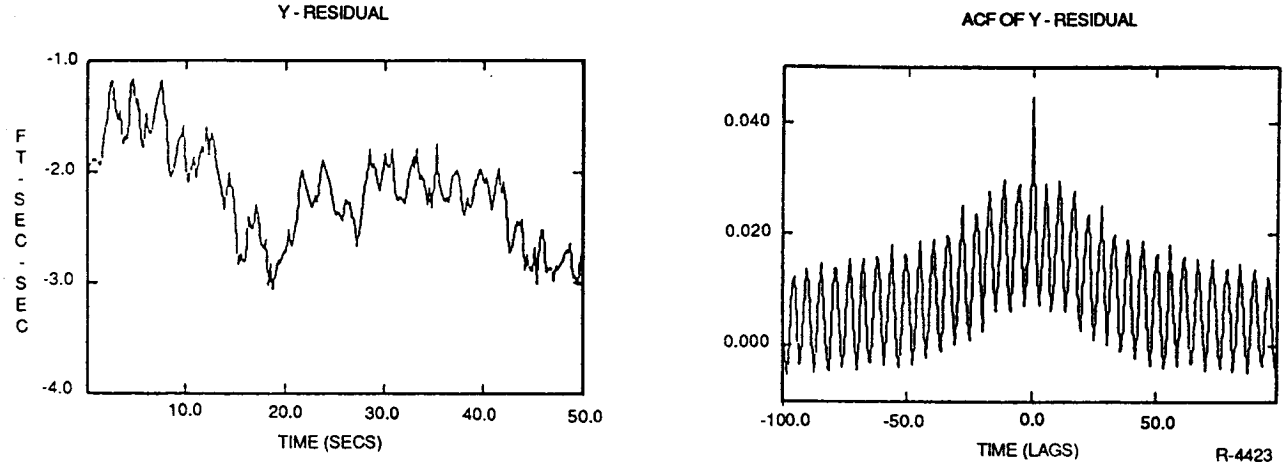


Figure 5-4. Translational Y-Residual and ACF

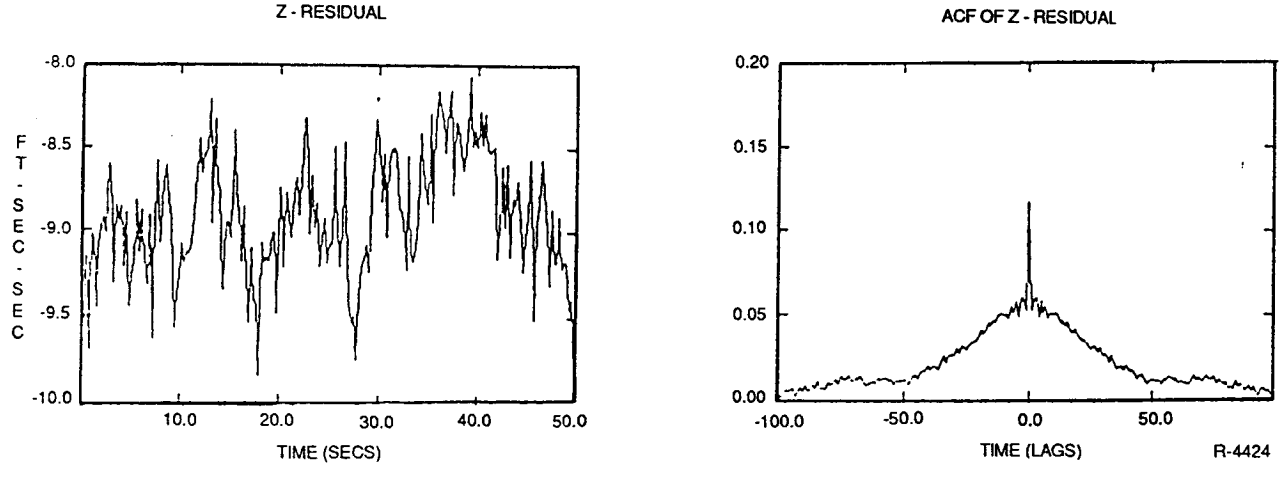


Figure 5-5. Translational Z-Residual and ACF

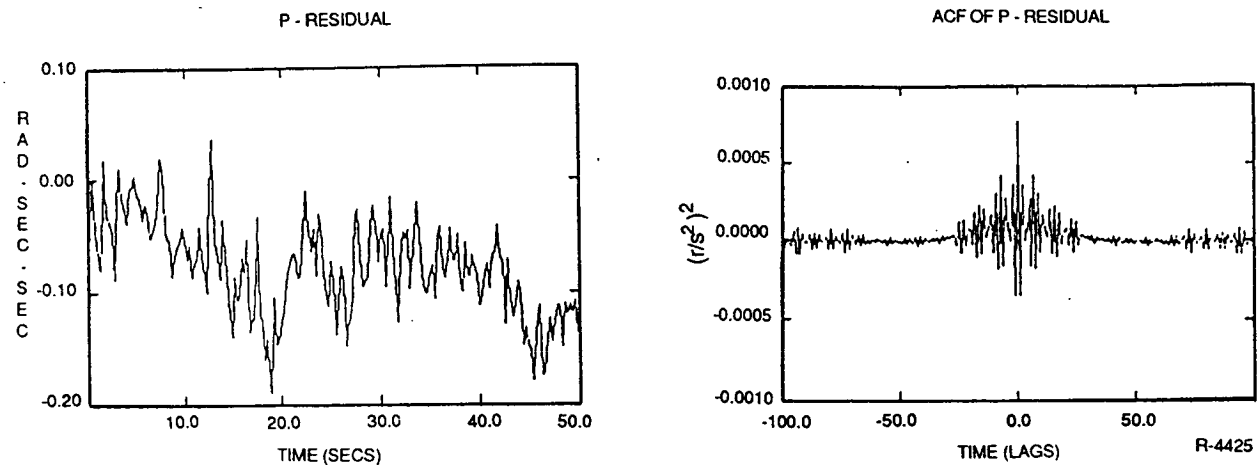


Figure 5-6. Rotational P-Residual and ACF

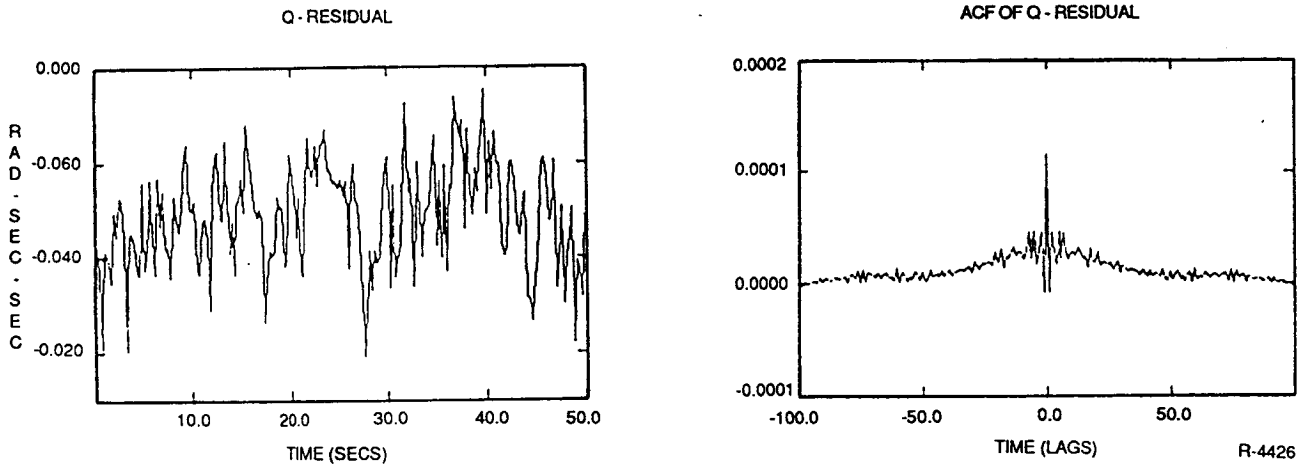


Figure 5-7. Rotational Q-Residual and ACF

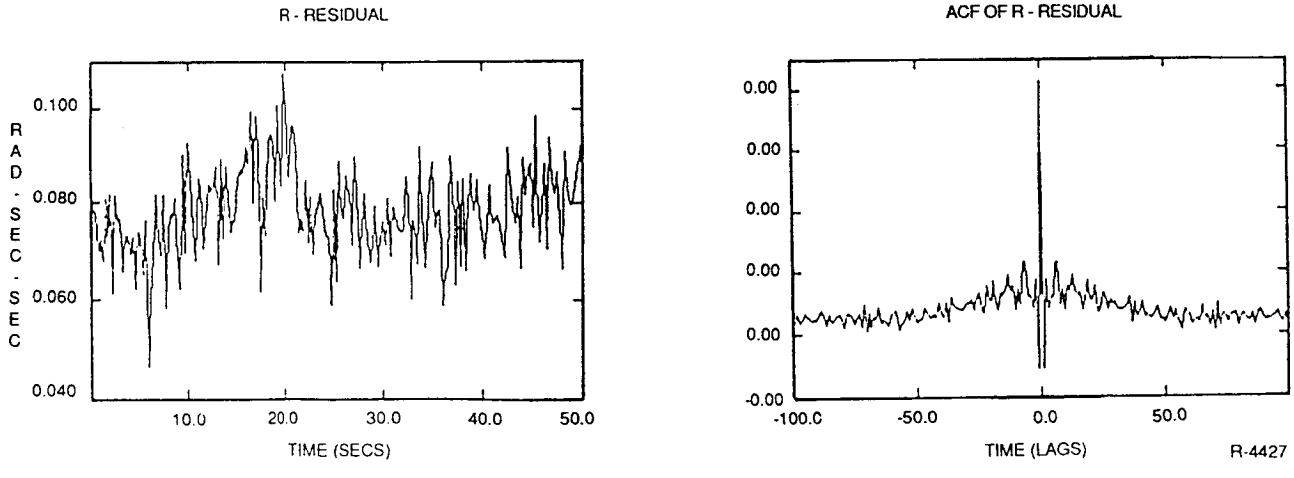


Figure 5-8. Rotational R-Residual and ACF

corresponding coefficients. Secondly, there is a nearly white component as evidenced by the near discontinuity at m=0. This component is due to electronic sensor noise and due to discretization error. The third error which is evident is a high frequency oscillatory error. The frequency of this error differs between residuals and may be due to vibration of the inertial platform. The oscillation frequencies vary from about 4 Hz to less than 10 Hz. Wing bending probably plays no part in these errors since the lower frequency oscillations are in the x and y residuals. In the roll residual, a damped oscillatory error at about 3 Hz may be due to wing bending. Finally, in many of the residuals, one or two first-order-like errors are evident. The time constants of these errors are around one second. These errors are almost certainly due to the excitation of parametric errors in the aerodynamic model by the (forced) motion of the aircraft.

POWER SPECTRAL ESTIMATION

The qualitative modal analysis discussed in the previous paragraph suggests that the estimation of the spectral density matrix ought to be accomplished using some of the more advanced time series methods (e.g., see [54]). However, time did not permit full investigation of these methods for the vector case. Therefore, spectral densities were estimated using the Blackmun-Tukey procedure (i.e., the discrete Fourier transform, DFT, of R(m)). That is,

$$S_n(\omega = \sum_{M=-100}^{100} R(m)\ e^{-jm\omega} \qquad (5\text{-}41)$$

for $\omega \in [0, 2\pi]$. Note that S_n is, in general, complex (except for the diagonal elements which must be real). Also we have the property $S_n(\omega+\pi) = S_n(\omega-\pi)^H$ where H denotes complex conjugate transpose. Figures 5-9 to 5-14 show the diagonal elements of S_n.

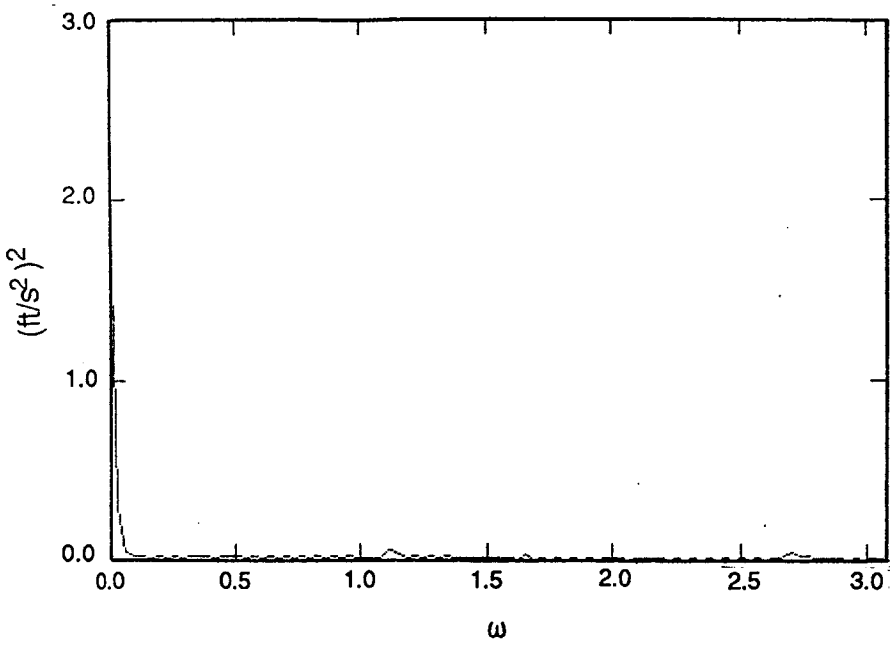


Figure 5-9. Spectrum of x Residual

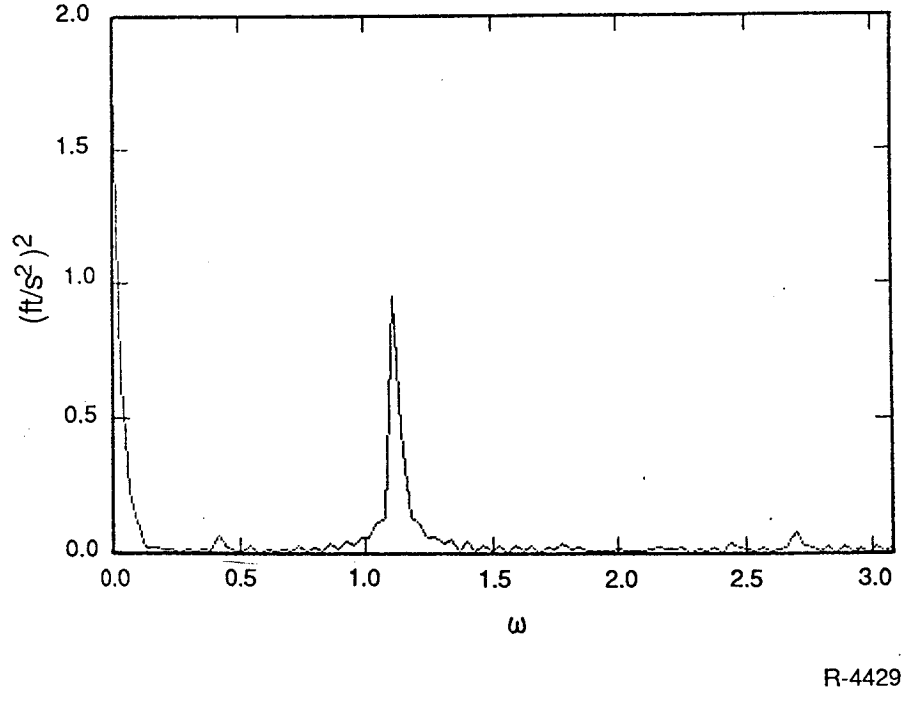


Figure 5-10. Spectrum of y Residual

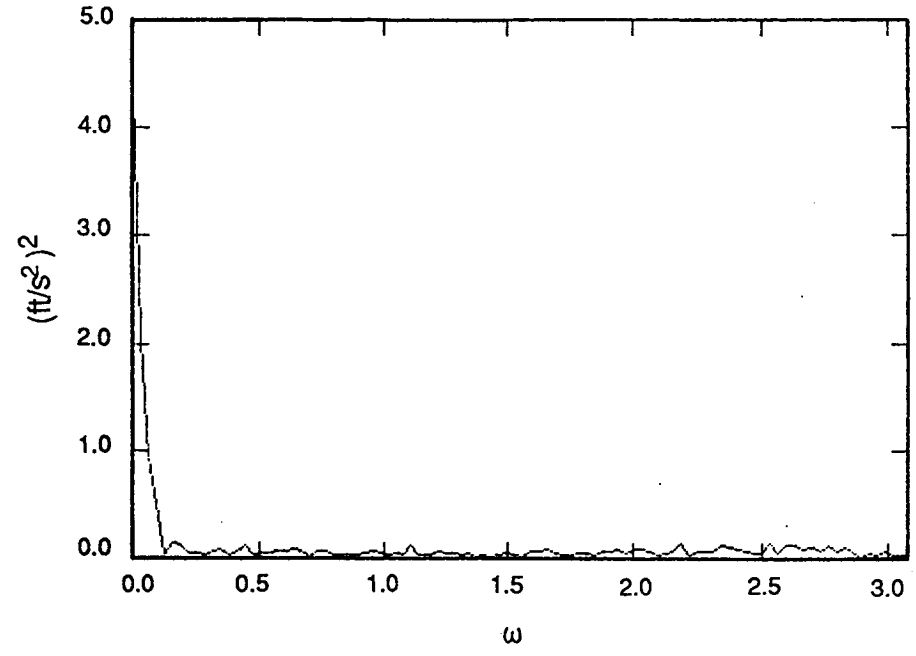


Figure 5-11. Spectrum of z Residual

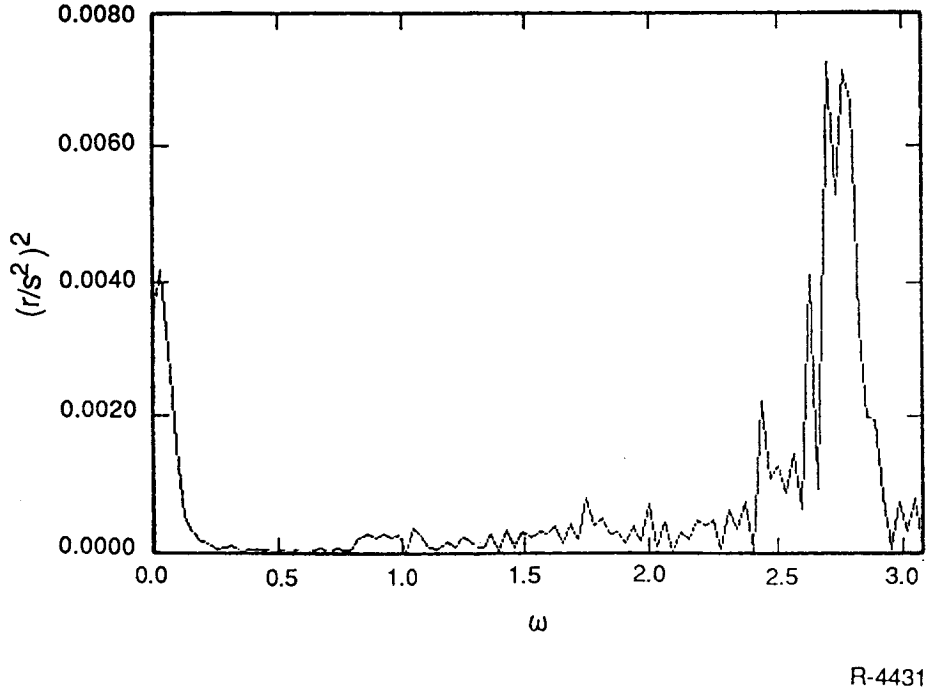


Figure 5-12. Spectrum of p Residual

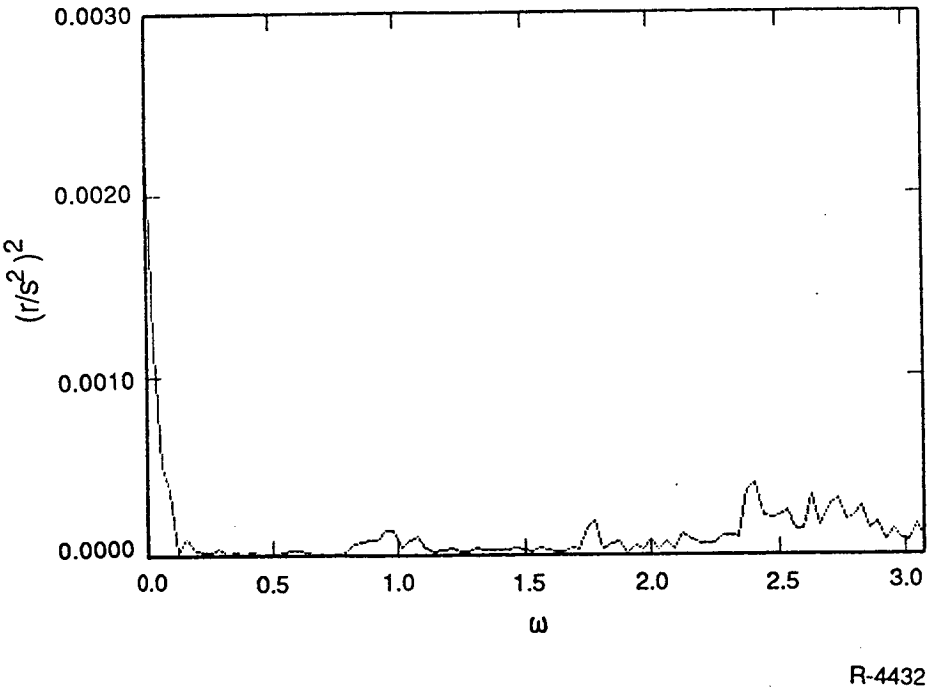


Figure 5-13. Spectrum of q Residual

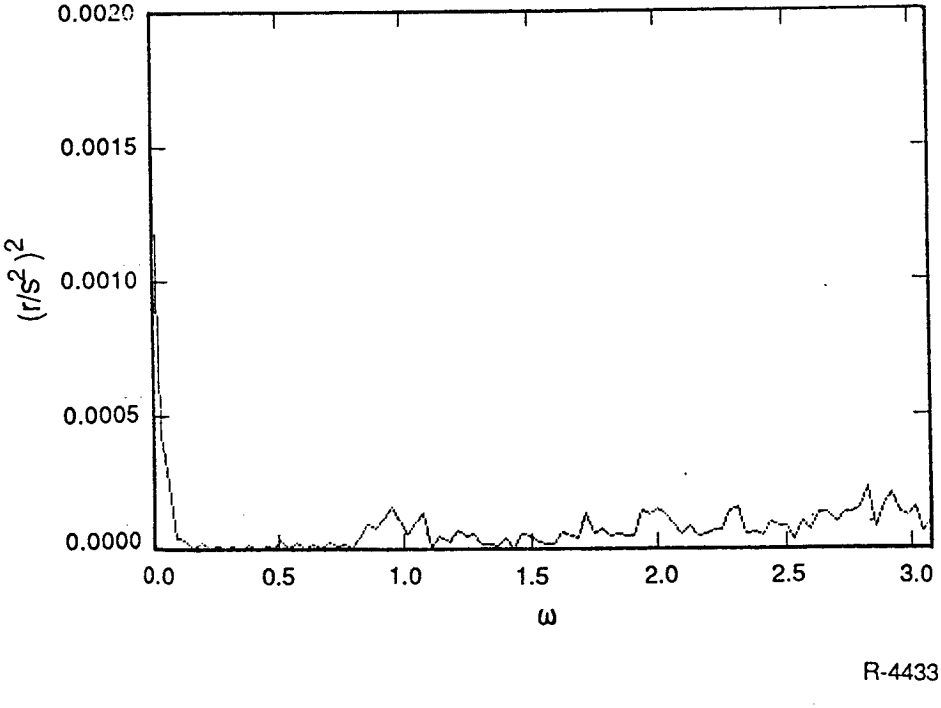


Figure 5-14. Spectrum of r Residual

The residuals we have generated typically have a large nonzero mean over any meaningful interval of interest. It is typical, however, to assume that this mean is zero when estimating (or modeling) power spectra. The tacit assumption is that for any interval of data, one can eliminate the mean value of the process and characterize only the deviations from that mean (assuming that the measured mean is actually the statistical mean). However, for our purposes we wish to characterize the low frequency energy which is indicated by the presence of the nonzero mean in the residuals. In the estimation of the autocorrelation function, it is typical to either remove the mean and then add in $\bar{\nu}\,\bar{\nu}T$ to R(m) or to perform what is known as a circular correlation [55] (i.e., assume that the observed sequence is periodic such that $\nu(k+N) = \nu(k)$). Both of these approaches address the problem that, for large nonzero mean processes, the computations given for R(m) in Eq. 5-40 result in a triangular autocorrelation sequence (which does not accurately reflect the, presumed, theoretical sequence) and, that because we only have estimates of the autocorrelation over a finite window, the resulting power spectral estimates may have negative values near (and sometimes not so near) zero frequency when $\bar{\nu}$ is large.

To get around these problems and still characterize the very low frequency behavior of the residuals, we assume that the observed mean value really represents a single sample from a filtered low frequency noise sequence. That is, since

$$\bar{\nu} = \frac{1}{N} \sum_{j=1}^{N} \nu(j) \qquad (5\text{-}42)$$

we know that most of the high frequency energy in ν will not have an impact on $\bar{\nu}$ when N is large. If we assume that the only component of ν which can effect $\bar{\nu}$ has a spectrum of

$$S_{LF}(\theta) = G(1-a^2)/\ \{(1-a\ \cos(\theta))^2 + (a\ \sin(\theta))^2\} \qquad (5\text{-}43)$$

then one can show that $E\{\bar{\nu}^2\}$ is approximately equal to G as long as $-\ln(a) << 2\pi/N$. Now since we have only one sample of $\bar{\nu}$ (for each window over which we wish to estimate the power spectrum) it is difficult to compute a meaningful statistical average. Therefore, we will assume that the observed value of $\bar{\nu}$ represents a typical value and model the low frequency energy in each residual using Eq. 5-43 with $G = \bar{\nu}^2$. In practice, we have chosen $-\ln(a) = 0.6/N$. Note Note that with typical values of ν and N, $\bar{S}_{LF}(0)$ is very large; however, this is consistent with the expectation that for very low frequency noise, signals with very large total energy will be needed to achieve detectability. The spectra $S_4(\theta)$ are then added to the spectrum in Figs. 5-8 to 5-13 to evaluate detectability.

FAILURE DIRECTIONS

The failure directions, $\hat{C}_1$, of Eq. 5-38 are given by the dimensional derivatives of each control element. Defining the six dimensional residual vector as $\nu^T = (\nu_x, \nu_y, \nu_z, \nu_p, \nu_q, \nu_r)$, failure directions are given by,

	x	y	z	p	q	r
Left Stabilator	.21 E-1	.0	-.25	.12 E-1	-.36 E-1	.88 E-3
Right Stabilator	.21 E-1	.0	-.25	-.12 E-1	-.36 E-1	-.88 E-3
Rudder	.0	.20	.0	.15 E-1	.0	-.17 E-1
Left Elevator	.99 E-2	.0	-.12	.56 E-2	-.17 E-1	.51 E-3
Right Elevator	.99 E-2	.0	-.12	-.56 E-2	-.17 E-1	-.51 E-3
Left Aileron	.11 E-1	.18 E-2	-.13	.12 E-1	-.47 E-2	.96 E-3
Right Aileron	.11 E-1	-.18 E-2	-.13	-.12 E-1	-.47 E-2	-.96 E-3

where the numbers are given in acceleration units per degree of deflection.

DETECTABILITY EVALUATION

To evaluate detectability, we considered two types of failure signatures: broad-band and narrow band. For narrow-band failure signatures, we can use Eq. 4-33 to determine the minimum signature energy needed to achieve some desired value of B^W for various frequencies of excitation. Figures 5-15 through 5-21 plot E_{min} versus frequency for each control element using $B_d = 3$ and the power spectral density estimates described above. As expected, large amounts of energy at low frequencies are needed because of the large mean values for the residuals. The plots also indicate that aileron and elevator failures may need large signature energies at high frequencies. This is probably due to the high frequency errors evidenced in the roll residual (ν_p).

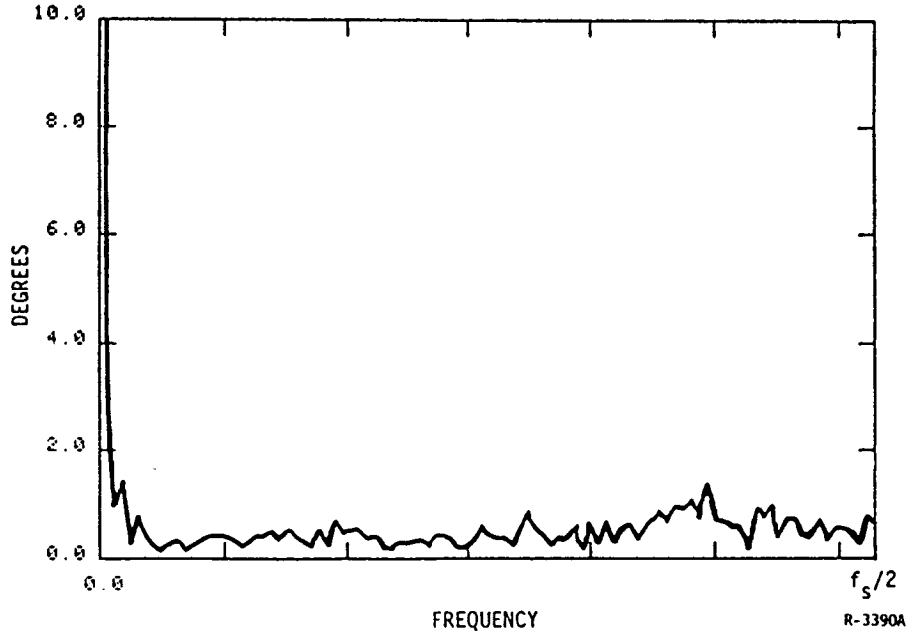


Figure 5-15. Detectability of Narrow Band, Left Stabilator Failure Signatures Versus Frequency

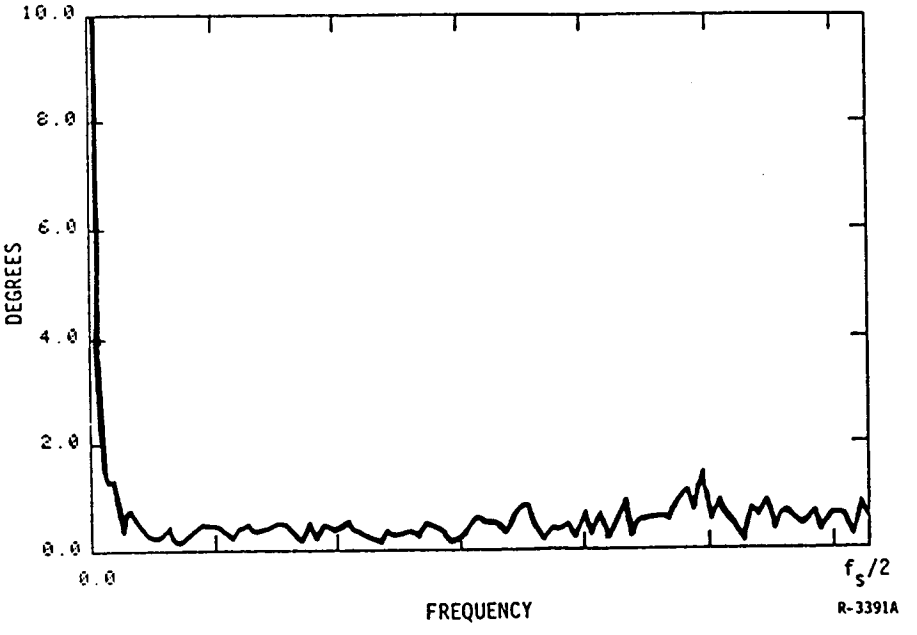


Figure 5-16. Detectability of Narrow Band, Right Stabilator Failure Signatures Versus Frequency

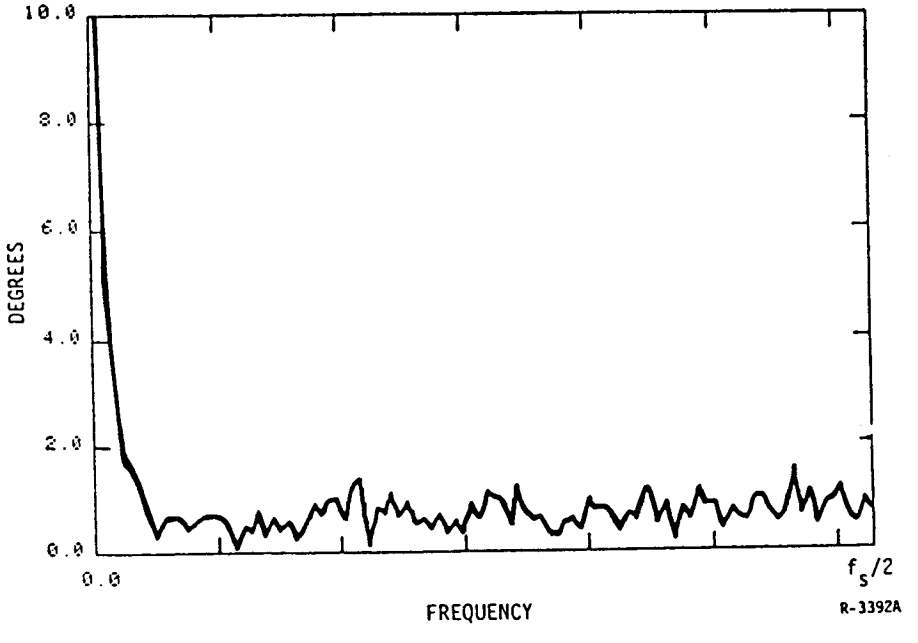


Figure 5-17. Detectability of Narrow Band, Rudder Failure Signatures Versus Frequency

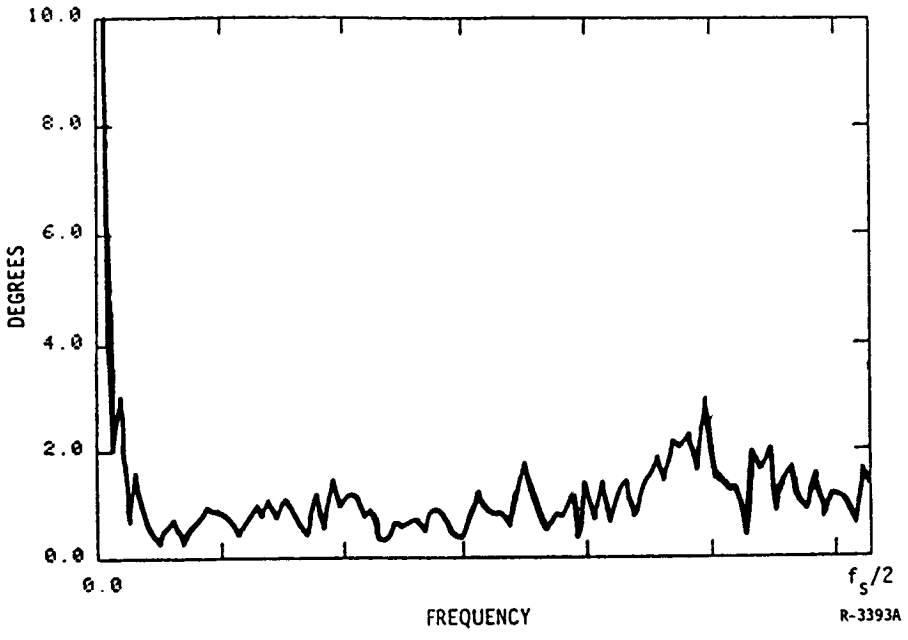


Figure 5-18. Detectability of Narrow Band, Left Elevator Failure Signatures Versus Frequency

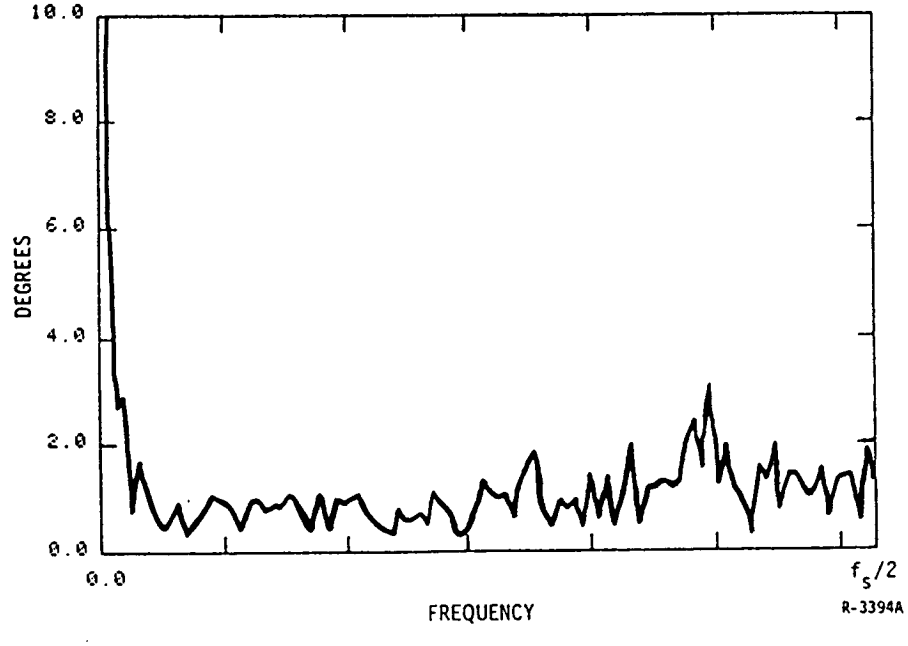


Figure 5-19. Detectability of Narrow Band, Right Elevator Failure Signatures Versus Frequency

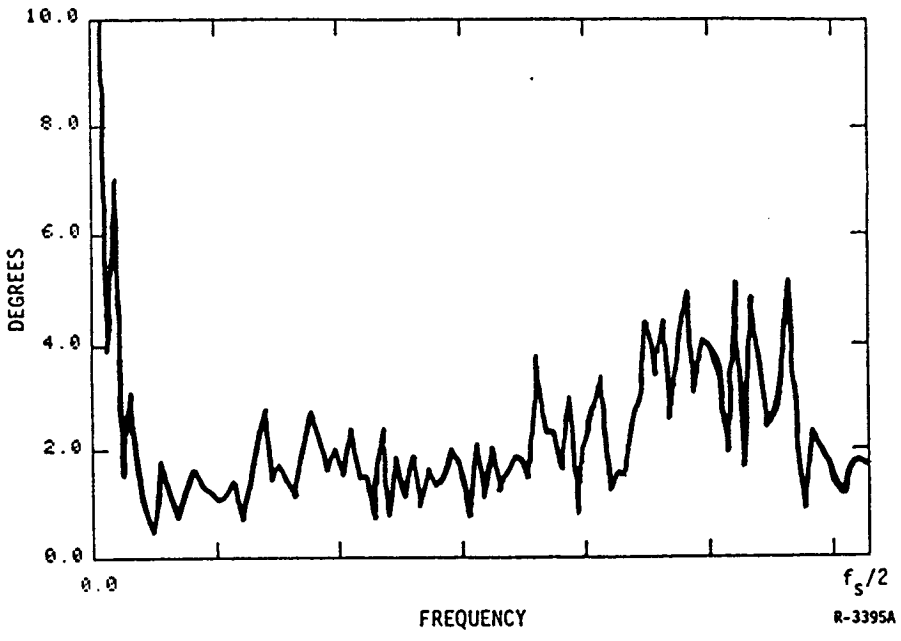


Figure 5-20. Detectability of Narrow Band, Left Aileron Failure Signatures Versus Frequency

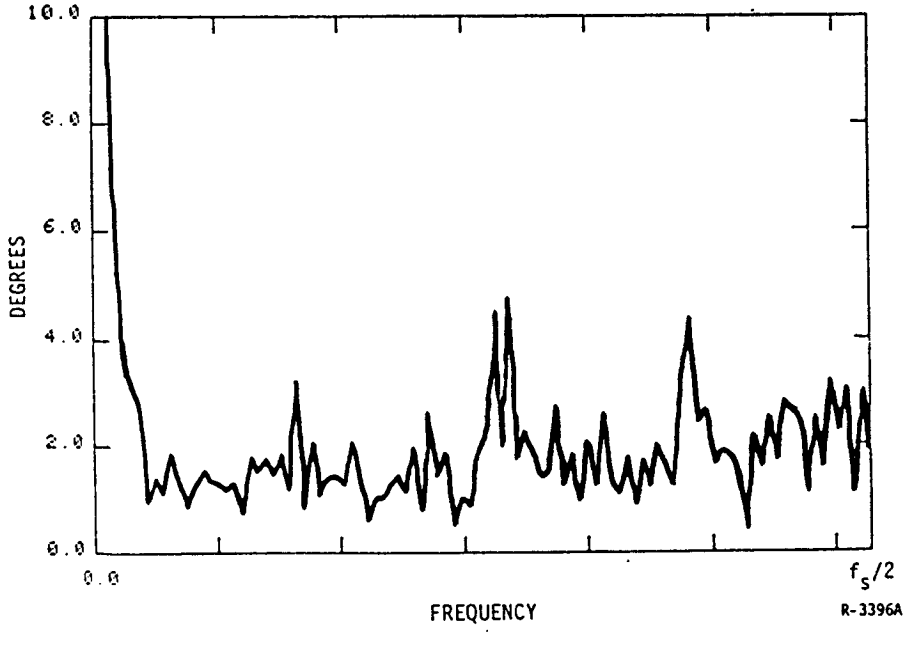


Figure 5-21. Detectability of Narrow Band, Right Aileron Failure Signatures Versus Frequency

To examine broadband failures signatures, we assume that the failure signature could be represented by,

$$m_j(\omega) = \frac{(1-a^2)}{(1-a\cos\omega)^2 + (a\sin\omega)^2} \qquad (5\text{-}44)$$

By using Eq. 5-44 in Eq. 4-33 and performing the required integrations numerically, we can determine the minimum broadband signature energies required for detection. Table 5-2 shows the results with a = .4 (~ 8 r/s bandwidth) and B^W_d = 6. Notice that the results are not symmetrical with respect to surfaces on either side of the airplane. This is most likely due to the fact that the errors which give rise to our power spectral estimates were not symmetrically excited during the interval of data we used.

TABLE 5-2. DETECTABILITY OF BROADBAND FAILURE SIGNATURES

Control Element	Total Signature Energy (degrees)
Left Stabilator	0.44
Right Stabilator	0.45
Rudder	0.62
Left Elevator	0.95
Right Elevator	0.98
Left Aileron	1.85
Right Aileron	1.92

The values in Table 5-2 are notably small since they represent total signature energies (i.e., the integral of the squared signature over time). Even signatures whose total energies are factor of 10 greater than those in Table 5-2 would be acceptable. A signature energy of 10 degrees can be achieved by a constant deflection of 0.5 degrees for one second at the 20 Hz sample rate. However, these numbers should only be treated as lower limits to performance (best case) for a very specific failure signature which is known and for a

perfectly known power spectrum. A more in-depth analysis might include pulse-like signatures since they may be more representative of actual signatures and might utilize PSD estimate from a broader set of flight data. Detectability in terms of pulse height versus pulse width could also be plotted. Furthermore, distinguishability of failures should be determined using the methods described in subsection 4.2.2.

5.1.3 Aircraft Path Decision Process Design

The goals and assumptions to be used in this design were detailed in Section 2. These included:

1. No multiple failures
2. Single flight condition
3. No additional aerodynamic failure effects
4. No detailed assumptions about failure signature models
5. Validated sensors

The control elements we will consider as possible failures in this design are left and right engines (LT and RT), ailerons (LA and RA), horizontal tail (LHT and RHT), and rudder (R). The horizontal tail is a fictitious surface which represents both stabilizer and elevator surfaces. This is done because our preliminary evaluation of distinguishability indicated that same-side elevator and stabilizer surfaces are indistinguishable based on force and moment balance magnitudes along (i.e., without using temporal signature information). Aircraft path engine failures are those occurring outboard of the "engine actuator" output measurement. Since we have chosen (Section 5.2) the engine actuator path as the path from the throttle command to the engine pressure ratio (EPR), aircraft path engine failures are those in which EPR follows the throttle command but the resultant thrust is inconsistent with EPR (e.g., thrust reversers deployed).

ALGORITHM STRUCTURE

The overall structure of the decision process will include trigger, verify, and isolate test procedures as discussed in Section 3.1. This structure is chosen as a trade-off between computational complexity and performance in solving the unknown onset time problem. For the aircraft path control element FDI problem, this structure takes the form shown in Fig. 5-22. The decentralized residuals defined in subsection 5.1.1 are the only inputs to this decision process. All decisions are, therefore, based on the relative sizes and spectral characteristics of these six residuals. The trigger process is a system monitor which is used to indicate the possible presence of abnormal behavior. In order to maximize the sensitivity of the trigger process to individual control element failures, separate trigger tests are used for each element. That is, each test is designed so that if a detectable failure of control i occurs, then its corresponding trigger test will "pass." Note that the conclusion that control i is failed if its corresponding trigger passes is invalid since such a conclusion is based on the converse of the previous sentence. One consequence of this structure is that it is possible to have redundancy in the trigger tests. For example, if a single residual is the only source of information for several triggers, then the resulting trigger tests will be all identical.

The verify and isolate processes are initiated when a trigger test passes. The verify process is initiated by the trigger process and is used to reject false triggers. The verify process, therefore, plays a major role in achieving the desired overall false alarm rate. The isolate process runs in "parallel" to the verify process and is responsible for making decisions regarding the identity of the failed control element. In principle, only the implied

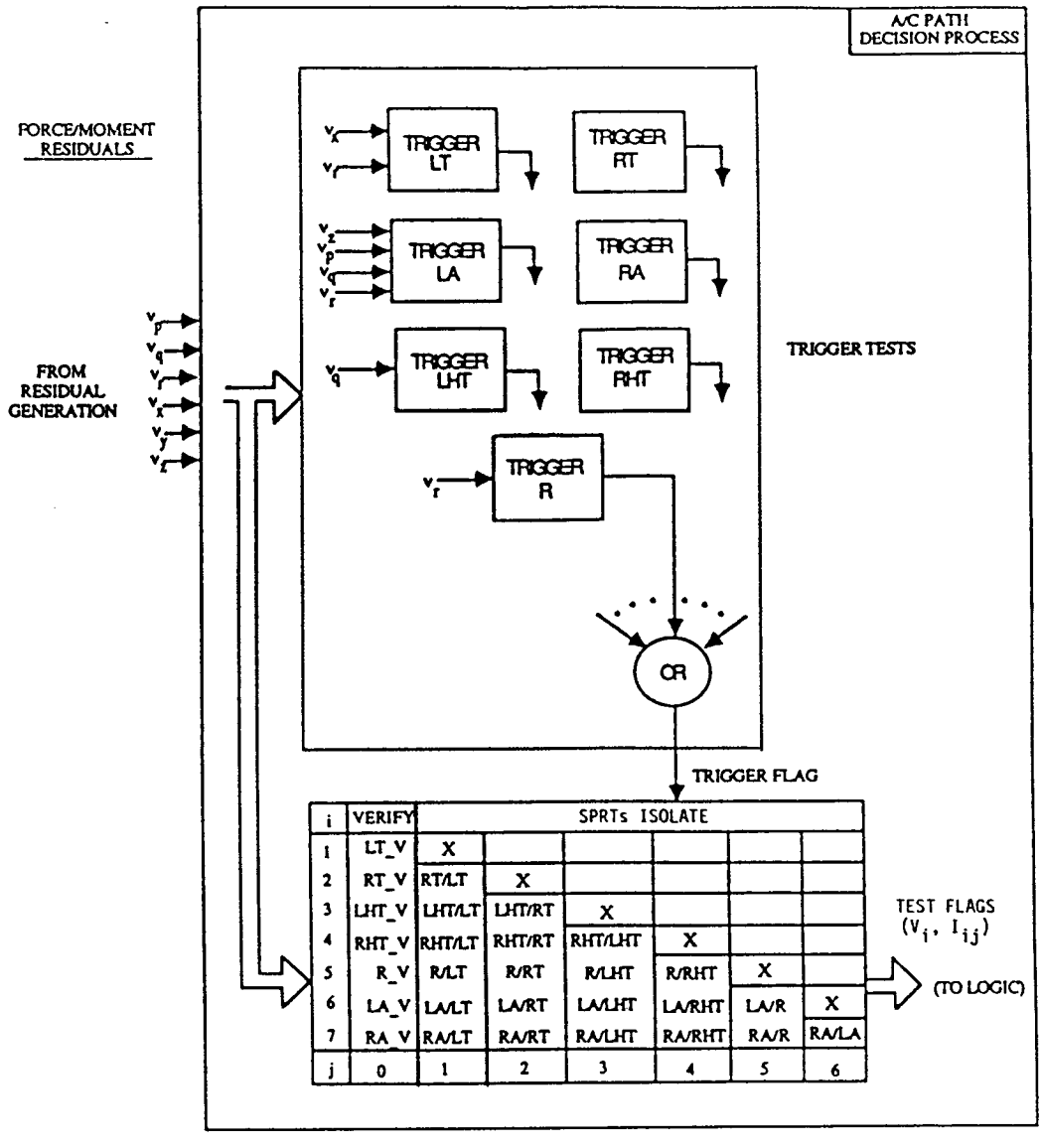


Figure 5-22. A/C Path Decision Process R-4434

ambiguity group resulting from a set of triggers (not the union of all trigger tests which pass) needs to be "isolated." However, for simplicity, all verify and isolate tests will be initiated following any passed trigger test for the current design.

The matrix of verify and isolate tests in Fig. 5-22 indicates that we will be using sequential tests which provide "votes" which indicate which of

any pair of system hypotheses is more likely (i=0 indicates normal operation; the control elements are indexed by i=1, ..., 7). The full matrix of isolation test flags, I_{ij}, take on values of 1 when j is more likely than i, 0 when i is more likely than j, and 2 when the sequential test has not yet completed. When the i,j^{th} test is completed and no decision can be made, I_{ij} and I_{ji} are both set to 0.

To declare that the j^{th} control element has failed, V_j=1 and I_{ij}=1 for all $j \neq i$. That is, the j^{th} column of the matrix in Fig. 5-22 must contain all 1's. Only one such column is possible due to the pairwise comparisons. Other information is available from this test matrix and some alternate uses are described in [56].

HYPOTHESIS TEST DESIGN DETAILS

The generic decision-process design procedure discussed in subsection 3.1 will be applied to this problem using the theoretical developments of subsection 3.3 The first step in this procedure is to develop structures for the hypothesis tests indicated in Fig. 5-22 based on a design model. For this model, we will assume that over the time interval needed to detect and identify failures, the failure signature is coherent (i.e., of one sign) and that this interval is substantially shorter than the time constants of some assumed very low frequency residual errors. We also assume (for defining test structures) that another residual error is broadband (i.e., white) noise associated with sensor noise and discretization errors. These errors will be modeled as Gaussian processes whose covariance characteristics are the same under all modes of operation.

After defining the hypothesis test structures from the above design model, the parameters of those tests must be chosen. To do this, we will hypothesize a statistical truth model for the residual error which is more detailed than the design model description. This model is then used to optimize parameters and perform sensitivity analyses. In general, such a model might take the form,

$$x_r(k) = A_r\, x_r(k-1) + B_r\, w_r(k-1) \tag{5-45}$$

$$\nu(k) = C_r\, x_r(k) + D_r\, w_r(k) \tag{5-46}$$

where $\nu(k)$ is the six dimensional residual, x_r is the n_r dimensional "truth-state" vector and w_r is a white noise vector with covariance matrix Q_r. In the design procedures to be discussed below, we will assume that the residuals are independent and that each residual can be represented by the sum of a white noise term, n_w, a first order low pass markov process, n_L, and a first order high pass Markov process, n_H. The white term represents sensor noise and discretizaton errors. The low pass term represents "in-band" errors (i.e., those errors which are excited in the same frequency range as the failure signatures). The high pass term represents unmodeled dynamics. Thus we have,

$$\nu_j(k) = n_w{}^j(k) + n_L{}^j(k) + n_H{}^j(k) \tag{5-47}$$

$$x_L{}^j(k) = a_L{}^j x_L{}^j(k-1) + (1-a_L{}^j) w_L(k-1) \tag{5-48}$$

$$x_H{}^j(k) = a_H{}^j x_H{}^j(k-1) + (1-a_H{}^j) w_H(k-1) \tag{5-49}$$

$$n_L{}^j(k) = x_L{}^j(k) \tag{5-50}$$

$$n_H{}^j(k) = w_H{}^j(k) - x_H{}^j(k) \tag{5-51}$$

The parameters of this model (cutoff frequencies and noise variances) are computed from an error budget which is described in subsection 5.1.4.

Prefilters - The use of colored, stochastic residual errors in the design model indicates that a "prewhitener" is necessary for each hypothesis test [32]. Ideally, this prewhitener is a multivariable filter which whitens and decorrelates the residual process (note that we don't really expect to be able to achieve a white process in reality since the residual errors are only modeled as stochastic; the true errors are, of course, nonstationary). While a multivariable filter may represent an optimal procedure for the design model, it, in general, has one major drawback for the current problem. In general, a multivariable filter can (depending on the nature of the frequency errors) take residuals which have failures that show up in fixed directions in residual space and transform them into signals which have failures that show up in time varying directions depending on the temporal characteristics of failure signatures. To retain the desired insensitivity to specific failure signatures, we will first project the residuals into a subspace which is appropriate for each test (to be based, in part, on the fixed failure directions) and then perform a one dimensional whitening operation. This structure will sacrifice performance in comparison to an optimal algorithm when the frequency characteristics of the residuals are substantially disparate, but guarantees insensitivity of every test to the detailed failure signature characteristics (note that results in the previous section indicate that residual errors are very similar and so little performance degradation is expected).

To choose the characteristics of the one dimensional "whitener" for each test, first note that the design model which includes low frequency and broadband errors implies that a filter with a high pass characteristic is needed.

Rather than hypothesize some detailed power spectrum and solve Weiner-Hopf equations [32] (or equivalently design Kalman filters for some detailed state space residual model), we will simply choose a high pass filter (HPF) structure and select its cutoff frequency. To do this, first note that the cutoff frequency must be higher than the bandwidth of the low frequency errors of the residuals (the very low frequency errors which vary with, e.g., velocity changes). Secondly, the time constant of the HPF must be larger than the FDI interval (so that the assumed coherent failure signatures are not washed-out). As a first cut, we will assume that all six residual errors are the same. The (very) low frequency error bandwidth is probably no higher than 0.1 r/s. The failure signatures are probably not coherent for more than about 2-5 seconds and the desired (longest) FDI interval is about two seconds which is consistent with the length of time that the signature is coherent. As a result, a first order high pass filter with cutoff frequency of 0.5 r/s is preliminarily chosen for each test. This frequency may change for tests which require longer FDI intervals (e.g., because the signatures may have only lower frequency content).

Triggers - The structure for the trigger test is based on the projection-HPF structure discussed above, and

1. the assumption that the projected/filtered residual error is white, and
2. the assumption that the projected/filtered failure signature is constant.

These assumptions are reasonable since we assumed that the failure signature is coherent and that the time constant of the high pass filter is much longer than the FDI interval. The resulting test takes the form,

$$S_T = \sum_{j=k}^{k-N_T+1} \gamma(j) \gtrless t_T \qquad (5\text{-}52)$$

which is just a moving window average of the projected and high-pass filtered residuals. This is shown in Fig. 5-23. The absolute value function takes into account the uncertain failure sign (see subsection 3.3). As discussed in subsection 3.3.1, this test will reliably detect coherent failures of some minimal magnitude and all failures which are larger than this minimal value. The parameters of this test must now be designed. For these purposes, it is convenient to think of the HPF as coming before the projection operation in Fig. 5-23 since the first cut design has all HPFs (i.e., for each test) identical.

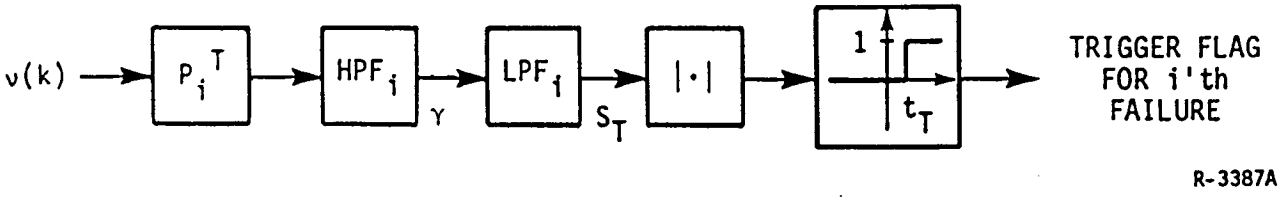


Figure 5-23. Structure of Trigger Tests - Time Variant

The parameters of the trigger test are the projection operator, the window length, N_T, and the trigger threshold. The selection of these parameters are based on the following concerns. The projection vector, P_i, must be maximally sensitive to the i^{th} failure direction while reducing the effect of real errors (i.e., those in a truth model) on S_T. Thus, the choice of P_i depends on the choice of N_T and a statistical model of the filtered residuals. The threshold, t_T, must be chosen to achieve a desired false trigger rate and it is, therefore, a function of P_i, N_T, and the statistical model. The window

length should be as short as possible to minimize trigger delay, but should be long enough so that important failures signatures will be detected. The specification of minimally detectable signatures is dependent on the choice of P_i, N_T, and t_T and the assumed statistical model.

Given Eqs. 5-47 through 5-51, the statistics of S_T can easily be computed as a function of P_i and N_T and the tradeoff between the false trigger likelihood and minimally detectable (bias) failure for various values of t_T determined. To select these parameters we proceed as follows.

1. Choose N_T as the longest allowable trigger time for the i^{th} failure.

2. Choose P_i to maximize the distinguishability of the i^{th} failure by maximizing the d^2 metric for the random variable S_T as a function of P_i assuming that the failure direction is C_i and that the HPFed signature is a constant. That is, compute the mean and variance of S_T under the hypotheses H_o and H_i from

$$E(S_T \mid H_o) = 0 \tag{5-53}$$

$$E(S_T \mid H_i) = N_T P_i{}^t C_i f_s \tag{5-54}$$

$$\text{Var}\ (S_T \mid H_o, H_1) = P_i{}^t \Sigma_f P_i \tag{5-55}$$

where Σ_f is the variance of $\left(\sum_{j=k}^{k-N_T+1} \nu(j)\right)$, and f_s is the magnitude of the constant failure signature (in units of the corresponding control element).* Using the results of subsection 3.3, we have

$$P_i = \Sigma_f^{-1} C_i \tag{5-56}$$

*Σ_f can be computed by forming a state variable equation for the (vector) filter corresponding to the desired sum and solving the discrete time Lyapunov equation [57]. This can be done by component by components when the components of $\nu(j)$ are uncorrelated.

3. Select the threshold to achieve a desired false trigger specification. For example, to achieve P_{FT} smaller than 10^{-3}, we need

$$t_T = 3\ \mathrm{Var}(S_T). \tag{5-57}$$

4. Find the value of f_s which achieves some specified miss probability. For example, to achieve P_{MT} smaller than 10^{-3}, we need

$$(E(S_T \mid H_1) - t_T) = 3\ \mathrm{Var}(S_T) \tag{5-58}$$

5. Determine which elements of the projection vector can be made equal to zero. This step is designed to enhance the robustness of the algorithm in the case where actual errors are larger than the truth model. The idea is to remove those residuals which contribute only marginally to detectability as measured by the size of f_s needed to achieve some P_{MT}. The procedure is iterative. The projection is restricted to have nonzero elements in only a subset of residuals and steps one to four are carried out. If f_s does not increase substantially (e.g., no more than one percent greater than its value with all residuals) then the residuals which are not included do not contribute to detectability and can be removed from that test.

The five steps just described provide a precise methodology for determining the values of the parameters of the trigger test in Fig. 5-23. Of course, these values depend highly on the statistical truth model assumptions, and, therefore, one must be careful in interpreting the results for nonstationary errors. For example, if Eqs. 5-48 to 5-52 are derived by considering observations of real errors during many different flying modes, a high degree of confidence in the resulting design may be allowed. In the remaining designing methodologies, we will also utilize Eqs. 5-48 to 5-52, and, therefore, the same comments apply.

Verify Test Design - The structure of the verify test is similar to the trigger test in that a "prewhitening" high-pass filter, projection operation occurs first. Assuming once again that the projected and filtered residual is zero mean and white during normal operation and nonzero mean with the same additive white noise during a failure, the optimal sequential probability ratio test takes the form,

$$S_V(k) = \left| \sum_{j=k_f}^{k} G_A Y(j) \right| - \sum_{j=k_f}^{k} G_D > t_V \quad \text{---> Verify passes}$$

$$< -t_V \quad \text{---> Verify fails} \qquad (5\text{-}59)$$

The test is started when a trigger occurs (i.e., at time k_f) and completes under the conditions in Eq. 5-59 or when a time limit, N_V, is reached. Figure 5-24 illustrates this test.

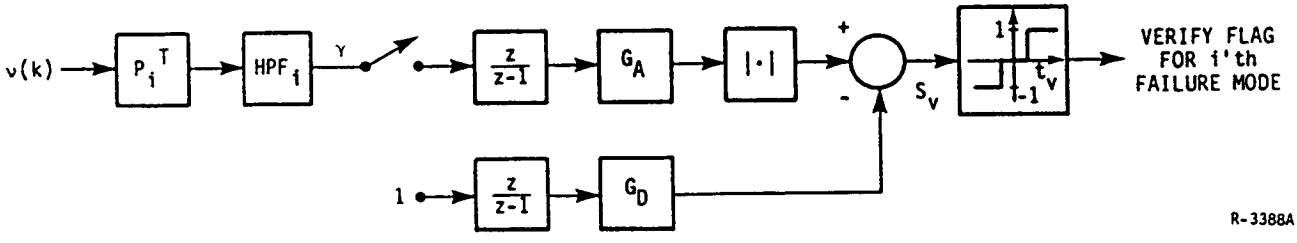


Figure 5-24. Structure of Sequential Verify Tests

The parameters of the verify test which must be chosen are P_i, G_A, G_D, t_V, and N_V. For a given value of t_V, G_D controls the speed at which false triggers are rejected and G_A controls the speed at which failures are detected. The relationship between G_A and G_D determines what failure sizes will tend to result in a passed verify test. As discussed in subsection 3.3.1, failure signatures greater than some minimally detectable value will be detected in shorter times by this test. As in the trigger test, P_i is chosen to maximize the sensitivity to failure i and minimize the effect of residual errors on the test. The tradeoffs which must be made in the design of a sequential test are similar to those in fixed sample size tests (P_{FA}, and P_{MD} versus failure size) except that we must include a probability of making no decision for tests which end in such a conclusion at their time limits.

Unfortunately, as we have indicated, the relationship between these probabilities and the test design parameters is very difficult to compute when there is a mismatch between the design and truth models. Thus, we must make use of heuristic methods for test design. One such method was discussed in subsection 3.3 and will be expanded here. The method proceeds as follows. Note that all expectatons are taken with respect to the pdfs defined by the truth model given in Eqs. 5-47 to 5-51.

1. Choose a maximum sample length N_V. Since the verify process is designed to validate triggers, $N_V = N_T$ is a reasonable choice.

2. Choose P_i to maximize the distinguishability of the i^{th} failure by maximizing the d^2 metric for the random variable $S_V(N_V)$ as a function of P_i assuming that the failure direction is C_i and that the HPFed signature is a constant. Note that unlike the trigger test we must specify a particular time at which d^2 is evaluated because the sequential test is a time varying test (i.e., it is triggered). The equations for choosing P_i are identical to those for the trigger test and if $N_V = N_T$, then the value of P_i is also the same.

3. Find the smallest value of $E(\gamma|H_i)$ which makes the fixed sample size test of length N_V reliable; where reliability is measured in terms of d^2 (large values of d^2 imply high reliability, see subseciton 3.3.1). Set G_A equal to this value and set $G_D = G_A^2/2$. This choice implies that failure signatures which result in average values of γ which are larger than $G_A/2$ tend to drive the sequential test to its positive threshold and those less than $G_A/2$ tend to drive it to its negative threshold.

4. Determine thresholds which result in low probabilities of false alarm, missing a minimally detectable failure, and making no decision. To accomplish this, consider the ideal SPRT. Let the desired detectable mean be m, and the process, γ, be white with variance σ^2. Then from [33] we have $G_A = m/\sigma^2$, $G_D = m/(2\sigma^2)$, and $t_V = -Ln[(P_M)/(1-P_F)]$, where P_M and P_F are upper bounds on the desired missed detection and false alarm probabilities. An interesting point in the test is when $E(S_V(k)|H_i)$ equals t_V (note that the symmetry of the test implies that this is the same point as when $E(S_V(k)|H_o)$ equals $-t_V$). Denote the value of k for which this is true by k_t. Then it is easy to show that the d^2 metric between the hypothesis of zero mean and the hypothesis of mean m (with additive white noise) for $S_V(k_t)$ is just $2t_V$. We now use this information to select thresholds for test Eq. 5-59 using the truth model distribution.

4a) Find k_t such that d^2 for $S_V(k_t) = -2Ln[(P_M)/(1-P_F)]$ (assume that the parameters of the test are given by steps two and three and that the hypothesis H_i includes a mean of minimally detectable magnitude; i.e., $E(\gamma|H_i) = G_A$.

4b) Set t_V to make $E(S_V(k_t)|H_i) = -E(S_V(k_t)|H_o) = t_V$, viz. $t_V = k_t\, G_A{}^2/2 = k_t\, G_D$.

5. As in the trigger test, an iterative procedure to zero projection elements which contribute little information to the test may follow. Note that this may be unnecessary if $N_V = N_T$ since the projections are then the same as those of the trigger test.

Isolation Test Design - The isolation tests are also started when a trigger occurs and completed when either a decision or a time limit, N_I, is reached. The isolation tests must have the property that failures which are larger than some minimal value should result in faster decisions. To achieve this we will use the failure magnitude invariant tests described in example three of subsection 3.3.2. In these tests, the projection operation we need to specify is orthogonal to one failure and maximally sensitive to another. Ideally, two "rejection" tests for each hypothesis pair would be implemented. However, to avoid some computational complexity, we will only perform one test for each hypothesis pair. As for the verify tests, the isolate test structure is based on the SPRT for detecting a constant in the projected and filtered residuals with additive white noise. The structure is shown in Fig. 5-25 and takes the form,

$$S_I(k) = \left|\sum_{t=k_f}^{k} \gamma_i(t)\right| - \left|\sum_{t=k_f}^{k} \gamma_j(t)\right| > t_I \longrightarrow \begin{array}{l}\text{test i over j passes}\\ \text{test j over i fails}\end{array} \qquad (5\text{-}60a)$$

$$< -t_I \longrightarrow \begin{array}{l}\text{test j over i passes}\\ \text{test i over j fails}\end{array} \qquad (5\text{-}60b)$$

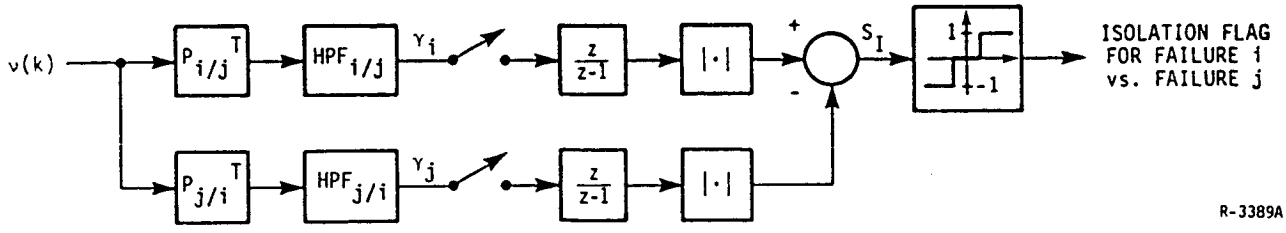


Figure 5-25. Structure of Sequential Isolation Tests

The parameters of the isolation test are the two projection operations $P_{i/j}$ and $P_{j/i}$ (both assumed to be unit vectors), t_I, and the time limit N_I. The projections are unit vectors because choosing otherwise would bias the test towards detecting smaller mean values in either γ_i or γ_j. (Note, such a case may be desirable since the average values of the corresponding control element which achieve equal size failures in γ_i and γ_j could be different). The projection $P_{i/j}$ is orthogonal to failure direction j and sensitive to failure direction i, and conversely for $P_{j/i}$. The threshold, t_I, and the time limit, N_I, control the error likelihoods and the probability of making no decision (since "no decision" will be the conclusion if neither Eq. 5-60a nor 5-60b is true when $k=k_f+N_I$). The sequential nature of this test, again, makes the relationship between the test parameters and the test goals difficult to compute. Therefore, a heuristic procedure similar to that used in the verify test is used. The procedure is described as follows.

1. Choose a maximum sample length N_I.
2. Choose the projection $P_{i/j}$ to maximize the d^2 metric for the quantity $\left(\sum_{t=k_f}^{k_f+N_I} \gamma_i(t) \right)$ subject to the constraint $P_{i/j}^t C_j = 0$, and similarly for $P_{j/i}$.

3. Find the smallest values of $\bar{s} = E(\gamma_i) = E(\gamma_j)$ which make the fixed sample size test of length N_I based on $S_I(N_I)$ reliable where reliability is measured in terms of d^2. To compute d^2 for $S_I(N_I)$ we have,

$$E(S_I|H_o) = N_I\bar{s} \tag{5-61}$$

$$E(S_I|H_j) = -N_I\bar{s} \tag{5-62}$$

$$Var(S_I|H_i, H_j) \leq (P^t_{i/j}\ \Sigma_f\ P_{i/j}) + (P_{j/i}\ \Sigma_f\ P_{j/i}) \tag{5-63}$$

where Σ_f is the covariance of $\sum_{t=k_f}^{k_f+N_I} \nu(t)$, where $\nu(t)$ is the HPFed residual vector, and where the inequality is due to neglecting the covariance between the absolute value terms in Eq. 5-60 (note that this inequality may give rise to some conservatism in the design of the isolation tests. A tighter bound would substract $2P_{i/j}\Sigma_f P_{j/i}$ from 5-13).

4. Find the value k_t at which d^2, for $S_I(k_t)$, is equal to $-2\ Ln[P_e/(1-P_e)]$, (P_e is a lower bound on the desired probability of error).

   Set $t_I = k_t\bar{s}$ (is $E(S_I(k_t)|H_i$ or $H_j) = t_I$ for failures which make $E(\gamma_i|H_i)$ or $E(\gamma_j|H)$ equal to $\bar{s}$.

5. Iterate on steps two through four to determine which residuals contribute little information to decision reliability. This is done by examining the increase in the average values of the <u>failure signature</u> required to achieve $\bar{s}$ for both failure modes. That is, let

$$f_i = \bar{s} \ / \ (P^t_{i/j}\ C_i) \tag{5-64a}$$

$$f_j = \bar{s} \ / \ (P^t_{j/i}\ C_j) \tag{5-64b}$$

A residual may be removed from the test (its projection zeroed) if, for example, the resulting value of $(f_i + f_j)$ is no bigger than 1.1 times itσ value with all residuals considered.

5.1.4 Detailed Design and Test Results

In this subsection, we present a detailed design of an aircraft path control element FDI system using the design procedures discussed in subsecion 5.1.3, evaluate its expected performance and demonstrate its capabilities using simulation data from NASA's nonlinear B-737 simulation. This subsection starts with a discussion of error budgets for use in the statistical truth model of Eqs. 5-47 to 5-51. Preliminary detectability results based on this error budget are computed by designing trigger filters for each of the seven control element failures considered as potential failures (left and right engines, ailerons and horizontal tails, and the rudder). Horizontal tail is a fictitious element used to represent elevator or stabilizer since these surfaces are indistinguishable. Several design iterations are then discussed qualitatively ending with the final error budget and design figures. Results of testing the resulting FDI algorithm using data from NASA's simulation are then presented. These results include false alarm checks and detection/isolation checks for various maneuvers and failures and a short investigation of alternative algorithm structures and threshold schedules. The subsection then concludes with a discussion of the results and conclusions about the FDI algorithm and the design procedure.

ERROR BUDGET ANALYSIS

For the design procedure described in subsection 5.1.3, a statistical truth model taking the form of three independent error sources for each residual was required. The three error sources consisted of a white noise term, a first order low-frequency term and a first order high-frequency term. An error analysis is now used to determine the bandwidths and energies in these processes.

All of the error analyses to be used here are based on a linearization of the residual equations. Table 5-3 shows the dimensional derivatives which comprise this model. Constant terms and velocity terms are neglected because they only contribute to very low frequency errors which are washed out by the "prewhitening" high pass filter. Both elevator and stabilizer surfaces are listed separately since they may not always be used proportionally in the control law, and spoiler panels are also listed since they may also be used.

TABLE 5-3. DIMENSIONAL DERIVATIVES

	α(deg)	β(deg/s)	p(deg/s)	q(deg/s)	r(deg/s)	TL/TR(klbs)	SL/SR(deg)	EL/ER(deg)	AL/AR(deg)	R(deg)	SPL/SPR
v_x (ft/s^2)	.576	0	0	.000309	0	.342	.0211	.00989	.0113	0	-.0181
v_y (ft/s^2)		- .663	.0131		.0372				±.00179	.202	±.0185
v_z (ft/s^2)	-3.60			-.00355		-.00474	-.248	-.116	-.133	0	.133
v_p (r/s^2)		- .0829	- .0347		.0124	±.0021	±.0119	±.00556	±.0118	.0147	∓.013
v_q (r/s^2)	- .0327			-.0113		.00620	-.036	-.0167	-.00473	0	.00197
v_r (r/s^2)		.0159	- .00267		- .00300	±.0124	±.000883	±.000506	±.000959	- .0174	∓.002

White Noise Budget - The white noise term represents errors due to sensor noise and discretization of the sensor measurements. The total variance of the white noise term is computed by taking the root sum square (rss) of the contributions of each measurement to each residual. That is,

$$\sigma^2_{n_w i} = \sum_j C^2_{ij} \sigma^2_j \qquad (5\text{-}65)$$

where C_{ij} represents the sensitivity of the i^{th} residual to the j^{th} measurement and σ_j is the standard deviation of the corresponding measurement noise. (Note that accelerometer measurements and angular acceleration computations are also taken into account in Eq. 5-65). Table 5-4 shows the sensor noise values used in this study.

TABLE 5-4. SENSOR NOISE BUDGET

Measurement	σ	Units
α	.4	degrees
β	.4	deg
p	.02	deg/s
q	.02	deg/s
r	.02	deg/s
TL/TR	.06	klbs
SL/SR	.1	deg
R	.1	deg
EL/ER	.1	deg
AL/AR	.1	deg
SPL/SPR	.1	deg
A_x	.32	ft/s/s
A_y	.32	ft/s/s
A_z	.32	ft/s/s

The resulting white noise budget (i.e., the $\sigma^2_{n_w}i$ in Eq. 5-65 with the C_{ij}s in Table 5-3 and σ_j^2 of Table 5-4) is shown in Table 5-5.

TABLE 5-5. WHITE NOISE BUDGET

Measurement	σ	Units
X	.39	ft/s/s
Y	.42	ft/s/s
Z	1.5	ft/s/s
P	.035	r/s/s
Q	.017	r/s/s
R	.012	r/s/s

Low Frequency (LF) Budget -The "low" frequency errors to be modeled are mostly due to parameter errors; i.e., errors in C_{ij} of Eq. 5-65. These errors are modulated by changes in the measurements y_j which multiply C_{ij} in the (linearized) residual generation equations. The first parameter to be specified for the LF errors is the cutoff frequency for each residual. The parameters $a_{LF}{}^i$ are then determined from

$$a_{LF}{}^i = e^{-\omega_{LF}{}^i(\Delta t)} \tag{5-66}$$

where $\omega_{LF}{}^i$ are the cutoff frequencies in radians/sec and Δt is the sample time. There are several ways in which the cutoff frequency could be computed. For example, examination of the closed-loop transfer functions (TFs) for the control law in use would indicate the frequency ranges in which each measurement may be excited. For the control law in [58], for example, transfer functions begin to roll off from anywhere between .2 r/s and 10 r/s. Such an

analysis indicates that more than a single LF error term may be needed since the different transfer function bandwidths can affect a single residual. In spite of this fact, the dominant TF roll off frequency is about 2 r/s. Since this is consistent with the covariance functions observed in subsection 5.1.2, we will use $\omega_{LF}^{i} = 2$ r/s for each residual.

The magnitude of the LF errors in each residual are specified in terms of the total variance of the process n_{LF}^{i}. Since these errors are mainly due to parametric errors, we must translate a specification of each parameter error into the total variance of the corresponding LF error process. There are several ways in which this can be accomplished, however, in every case, some notion of a design envelope is required.

The design envelope is introduced in order to determine the potential size of the contribution of any parametric error to the residuals. The design envelope amounts to a specification of the largest variations of each measurement in the bandwidth of interest. Since the HPF washes out errors below .5 r/s and the LF bandwidth is 2 r/s, the bandwidth of interest is [.5,2] r/s. Table 5-6 shows the design envelope used in this project. Note that we assumed that spoilers are not used in the control law (but are used in the residual generation process, thus affecting the white noise budget). These numbers were determined from representative simulations of the aircraft response to various maneuvers using the control law of [58].

The total variance of each LF term is now computed as follows. First specify the maximum variation (in percent) for each coefficient C_{ij}. Find the worst case error for each residual and then divide by three to get the total standard deviation of the corresponding LF term. That is,

TABLE 5-6. DESIGN ENVELOPE

Measurement	Max Variation	Units
α	5.0	degrees
β	5.0	deg
p	20.	deg/s
q	10.	deg/s
r	10.	deg/s
TL/TR	3.0	klbs
SL/SR	10.	deg
R	10.	deg
EL/ER	10.	deg
AL/AR	10.	deg
SPL/SPR	0.0	deg
A_x	10.	ft/s/s
A_y	2.0	ft/s/s
A_z	15.	ft/s/s

$$\sigma_\eta{}^i = \frac{1}{3} \sum_j \left| \Delta C_{ij}\, y_j{}^{max} \right| \qquad (5\text{-}67)$$

where ΔC_{ij} is the worst case error in C_{ij} and $y_j{}^{max}$ is the corresponding measurement's design envelope. The division by three is used because the standard deviation represents a typical value and we used a worst case analysis in summing errors. Furthermore, the design methodology detailed in 5.1.3 utilized measures in which error probabilities are roughly equal to a 3σ significance level. Thus, it is expected that the worst case LF errors will just barely cause decision "errors" (actual errors or no decisions) to be made in the hypothesis testing procedures. For five percent errors in every coefficient and the design envelope of Table 5-6, the LF residual errors are given in Table 5-7.

TABLE 5-7. PRELIMINARY LF ERROR BUDGET (1-σVALUES)

Measurement	σ	Units
X	0.096	ft/s/s
Y	0.10	ft/s/s
Z	0.47	ft/s/s
P	0.033	r/s/s
Q	0.024	r/s/s
R	0.0077	r/s/s

At this point it is worthwhile to compare the values in Table 5-7 to the control derivatives in Table 5-3 to get some idea of detectability. For example, we'd expect the pitch residual to play a major role in detecting elevator failures. However, in order for the effect of an elevator failure to be reliably detected, it should be about six times greater than the 1-σ value of the corresponding LF error ($d^2 = 6 \dashrightarrow P_e = 10^{-4}$). This would imply a need for a an average deflection of 8.6 degrees! Although smaller failures could be detected, this is the minimum which achieves the desired reliability. Thus, it appears that the design envelope may be too conservative. More will be said about this in subsequent discussions.

High Frequency (HF) Budget - High frequency errors in the residuals may be caused by neglected sensor dynamics and neglected flexure modes. For this study, we assumed that only the first-asymmetric and first-symmetric wing bending modes contribute to high frequency error. Furthermore, we assumed that these modes are excited above about 15 r/s.

The first order high pass Markov process described in subsection 5.1.3 is a poor model for the effects of these errors. This is true for two reasons.

First, the frequency shape does not correspond to the "peaky" nature of the real errors, and second, the power spectrum defined by the cutoff frequency of 15 r/s (with the sample period of .05 sec.) is virtually the same as any other spectrum with cutoff frequency above about 10 r/s. The latter effect is due to the specific realization of the HF errors which in the limit (as $\omega_{LF} \rightarrow \infty$), looks like a one time step lag. Due to these shortcomings, and to the lack of wing bending models in the simulation to be used for testing purposes, the final design will be based on no HF errors. Nevertheless, it is instructive to see an example of HF error budget modeling.

For the first asymmetric wing bending mode, the primary residual affected is ν_p. The dynamic errors associated with this mode are excited mostly by differential aileron motion and somewhat by rolling motion due to rudder deflection. Assuming that the major contribution is due to differential aileron, we can define the total variance in the HF error in ν_p by

$$\sigma_p^{HF} = \frac{1}{3} \left| C_{p\delta A} \, \Delta\delta_A^{max} \right|$$

$$= \frac{1}{3} (.012)(20) = 0.08 \qquad (5\text{-}68)$$

Similarly, for the first symmetric wing bending mode, the primary residuals which are affected are ν_q and ν_z. The HF errors in both residuals are excited by collective aileron deflection and by vertical wind gusts. Assuming that vertical wind gusts result in a change of α of no more than five degrees the total variance of the resulting HF errors is,

$$\sigma_q^{HF} = \frac{1}{3}\left[\left|C_{q\delta A}\ \Sigma\delta_A^{\max}\right| + \left|C_{q\alpha}\ \alpha^{\max}\right|\right]$$

$$= \frac{1}{3}\left[(.005)(20) + (.033)(5)\right] = .088 \qquad (5\text{-}69)$$

$$\sigma_z^{HF} = \frac{1}{3}\left[\left|C_{z\delta A}\ \Sigma\delta_A^{\max}\right| + \left|C_{z\alpha}\ \alpha^{\max}\right|\right]$$

$$= \frac{1}{3}\left[(.13)(20) + (3.6)(5)\right] = 6.9 \qquad (5\text{-}70)$$

DESIGN ITERATIONS

Several design iterations involving changes to the LF error specifications were made before a final specification was utilized in a full scale design. A short diary of these iterations is given here. In every design iteration it is assumed that the maximum length for a test is one second (N_{max} = 20) and that a first order HPF with cutoff frequency of 0.5 r/s is employed as a preprocessor.

The first design iteration utilized the error budgets developed above to design trigger tests. For the one second maximal trigger window length, projections were computed, false alarm thresholds were set to achieve (nominally) $P_F = 10^{-4}$ (i.e., at $3\sigma[S_T(20)]$), and the average values of control "deflection" required for reliable detection (i.e., $1-P_D = 10^{-4}$ for Gaussian statistics) were determined. This information is shown in Table 5-8. The projections are computed without any attempt to zero useless components (see subsection 5.1.3). Also, the aileron projections are incorrect due to a transcription error in the design software data. The detectability level is the average value of control deflection required for reliable detection (as defined in Table 5-8.

TABLE 5-8. AVERAGE VALUES OF CONTROL DEFLECTION REQUIRED TO ACHIEVE $1-P_D=10^{-4}$ (SIGNATURE "SIZE" 6σ [$S_T(20)$])

CONTROL		TL/TR	SL/SR	EL/ER	AL/AR	R
Projection (For left side where different)	X	.089	.013	.013	.017	0
	Y	0	0	0	.001	.020
	Z	0	-.006	-.006	0	0
	P	.012	.17	.17	.16	.061
	Q	.068	-.97	-.97	-.96	0
	R	.99	.16	.16	.21	-.99
DETECTABILITY LEVEL		1.97 klbs	3.2 deg	6.4 deg	6.9 deg	2.3 deg

All these values are quite large, as expected from preliminary comparisons of control derivatives with the LF error budget. It is deemed that these values are not satisfactory.

The only way that detectability levels can be reduced is by restricting the envelope of operation such that low frequency errors are reduced. This is obviously not a satisfactory solution by itself. However, the use of threshold scheduling permits us to reduce the nominal envelope and achieve the same false trigger alarm rates by scheduling thresholds during periods where errors are expected to be large. The resulting detectability levels are, of course, only valid for flight within the nominal envelope and deviations from this envelope will cause thresholds to be increased and detectability levels to increase.

We first considered eliminating the largest errors from the LF budget. This then implies that schedules which are a function of the corresponding measurements will be necessary. The largest errors are due to coefficients of:

1. α for ν_x
2. β for ν_y
3. α for ν_z
4. β and p for ν_p
5. SL/SR for ν_q
6. r and β for ν_r

Setting the errors in these coefficients equal to zero results in a smaller LF error budget and detectability levels of

TL/TR : 1.4 klbs

SL/SR : 1.7 deg

EL/ER : 3.4 deg

AL/AR : 3.6 deg

R : 1.5 deg

These detectability figures are more acceptable.

The last design iteration was accomplished after observation of the size of the residuals generated by processing simulation data from NASA's B-737 simulation. The observed residuals during a mild climbing turn maneuver were consistent with the maximum LF (model) errors used in the previous error budget. However, during roll, pitch and yaw doublets, were substantially (in magnitude) larger than what would have been predicted by this budget.

Since we ignored the largest errors in the previous budget, one explanation of the observations is that it is <u>these</u> errors which are being excited and should be accounted for by threshold scheduling. However, the temporal characteristics of the residuals are not correlated with the coefficient errors which were ignored in this budget. Thus, some other means of developing reasonable error budgets is necessary.

To expedite our further work, a final design was created using an error budget based on observed errors. Using 1-σ values corresponding to 2/3 of the maximum LF error observed during the climbing turn maneuver, we expect that errors which are twice as large as those observed during this maneuver to be adequately handled. The standard deviations of the LF errors used in the final design are,

x : .20 ft/s/s

y : .033 ft/s/s

z : .20 ft/s/s

p : .020 r/s/s

q : .0050 r/s/s

r : .0034 r/s/s

These budgets are larger than the previous design iteration for X, Y, Z, and P residuals and smaller for Q and R residuals. The resulting reliable detectability levels for the triggers (now with useless projection elements zeroed) are,

TL/TR : 1.8 klbs

SL/SR : 1.0 deg

EL/ER : 2.0 deg

AL/AR : 5.6 deg

R : 1.5 deg

Table 5-9 shows the "optimized" trigger projections. Notice that for detecting aileron failures, several residuals are required. This is due to the low overall effectiveness of the ailerons and to the fact that their effectiveness is spread amongst many axes. The use of only the r residual for rudder detection and only the q residual for horizontal tail detection is consistent with

TABLE 5-9. OPTIMIZED TRIGGER AND VERIFY PROJECTIONS FOR FINAL BUDGET

CONTROL		TL/TR	HORIZONTAL TAIL SL/SR & EL/ER	AL/AR	R
Projection	X	.011	0	0	0
± indicates sign for left/right controls	Y	0	0	0	0
	Z	0	0	-.006	0
	P	0	0	±.189	0
	Q	0	1.0	-.905	0
	R	±1.0	0	±.380	1.0

qualitative expectations, as is the use of both x and r residuals for aircraft path engine failures.

The procedures for designing verify and isolate tests described in subsection 5.1.3 were followed and projections, thresholds, and distinguishability levels computed. Table 5-10 summarizes the distinguishability results. Each entry in the table represents a particular isolation test. The first number corresponds to the average control deflection needed for the test to "pass" (pass means that the test decides in favor of the control identified by the column heading) and the second number corresponds to the average deflection needed for the test to "fail" (fail means that the test decides in favor of the control identified by the row heading). The largest of these numbers represents the overall distinguishability level for each surface and is shown at the right in Table 5-10. The tests corresponding to the largest values are highlighted.

Finally we note that the detectability and distinguishability levels are substantially larger than those determined in subsection 5.1.2. This is because the error budgets used in the design process were substantially larger

TABLE 5-10. DISTINGUISHABILITY LEVELS FOR ISOLATION TESTS

	RT	LHT	RHT	R	LA	RA	OVERALL DISTINGUISH-ABILITY LEVELS
LT	1.5 / 1.5	1.8 / .68	1.8 / .68	(1.9) / (1.7)	(1.9) / 3.5	1.8 / 3.9	1.9 klbs
RT	---	1.8 / .68	1.8 / .68	(1.9) / (1.7)	1.8 / 3.9	(1.9) / 3.5	1.9 klbs
LHT	---	---	(3.7) / (3.7)	.62 / 1.3	2.4 / (6.0)	1.2 / 5.5	3.7 deg
RHT	---	---	---	.62 / 1.3	1.2 / 5.5	2.4 / 6.0	3.7 deg
R	---	---	---	---	1.4 / 3.4	1.4 / 3.4	1.7 deg
LA	---	---	---	---	---	4.3 / 4.3	6.0 deg

than the observed values and because the FDI algorithm was purposely not optimized for the statistical (truth) model used in the evaluation process.

SIMULATION RESULTS

Simulations of a modified B-737 aircraft were performed at the NASA Langley Research Center and the required data were recorded on magnetic tape and delivered to ALPHATECH for use in simulating the FDI algorithm. The NASA simulation is a full six degree of freedom nonlinear simulation with nonlinear aerodynamic coefficients, accurate actuation models, sensor errors, and turbulence simulation using the Dryden spectra [59]. A total of 41, 60-second, simulation runs were made encompassing many categories of tests. For aircraft-path failures, these categories included:

1. False alarm checks with three "doublet" maneuvers and a climbing turn, with and without turbulence and sensor noise,

2. Detection checks for totally missing control surfaces, with and without turbulence and sensor noise, and

3. Detection checks for varying degrees of partially missing failures with and without turbulence and sensor noise.

In all of the runs, the control law defined in [58] was used.

The results to be presented next are for the climbing turn maneuver. Discussions of performance during the doublet maneuvers follows. Investigations of alternate algorithm structures (threshold scheduling and forced decisions) are presented next and we conclude with a discussion of the lessons learned.

Climbing Turn Results (No Failure) - Simulation data of the climbing turn maneuver with no failure were processed by the FDI algorithm. The climbing turn maneuver is accomplished by commanding pitch and bank angles (regulated by the control law of [58]). The commands are given by

At 10 sec $\dot{\theta}_c = 7.5$ deg/sec and $\dot{\phi}_c = 1.5$ deg/sec

At 14 sec $\dot{\theta}_c = \dot{\phi}_c = 0$

At 21 sec $\dot{\theta}_c = -7.5$ deg/sec and $\dot{\phi}_c = -1.5$ deg/sec

At 25 sec $\dot{\theta}_c = \dot{\phi}_c = 0$

Three cases were examined: no turbulence or sensor noise, sensor noise but no turbulence, and both turbulence and sensor noise present. During each of these runs, no false trigger was ever recorded.

The maneuver begins at ten seconds. Figure 5-26 shows the attitude, angular rates, and linear accelerations of the aircraft during the maneuver. Figures 5-27 through 5-32 show the six residuals for the turbulence and sensor

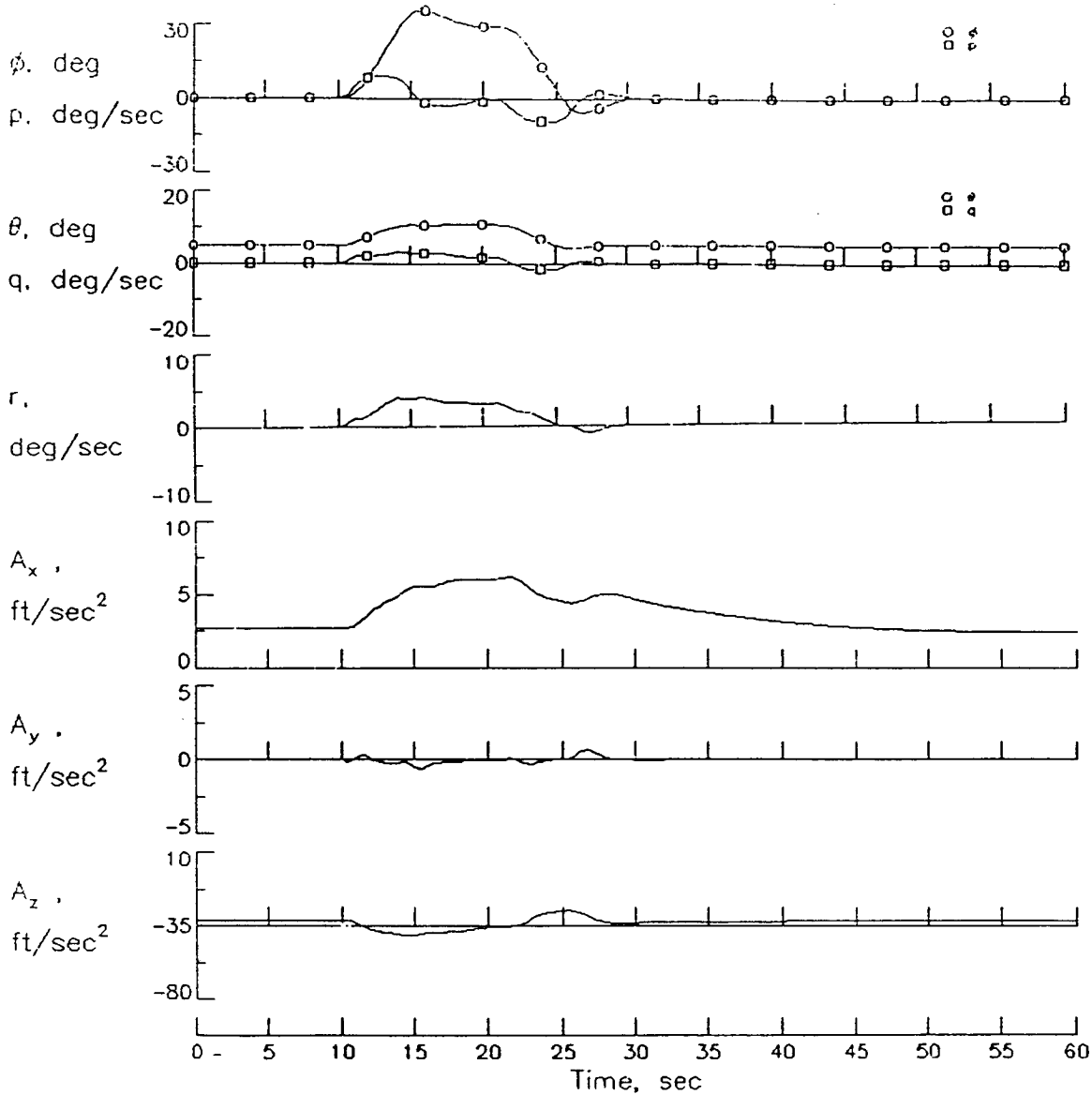


Figure 5-26. Normal Aircraft Response During Climbing Turns

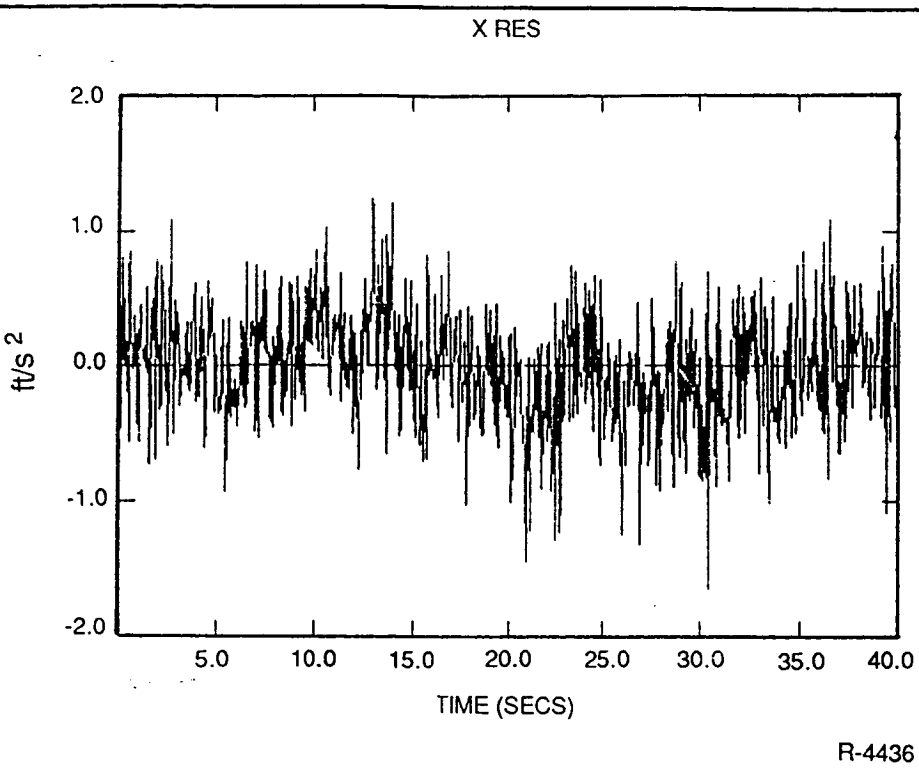


Figure 5-27. X RES

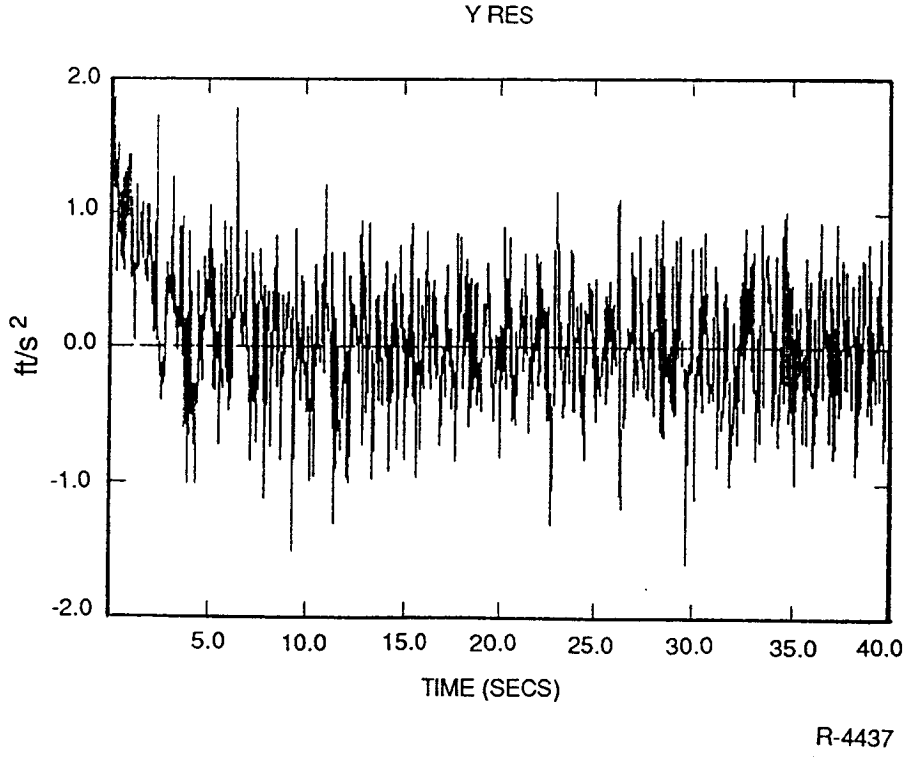


Figure 5-28. Y RES

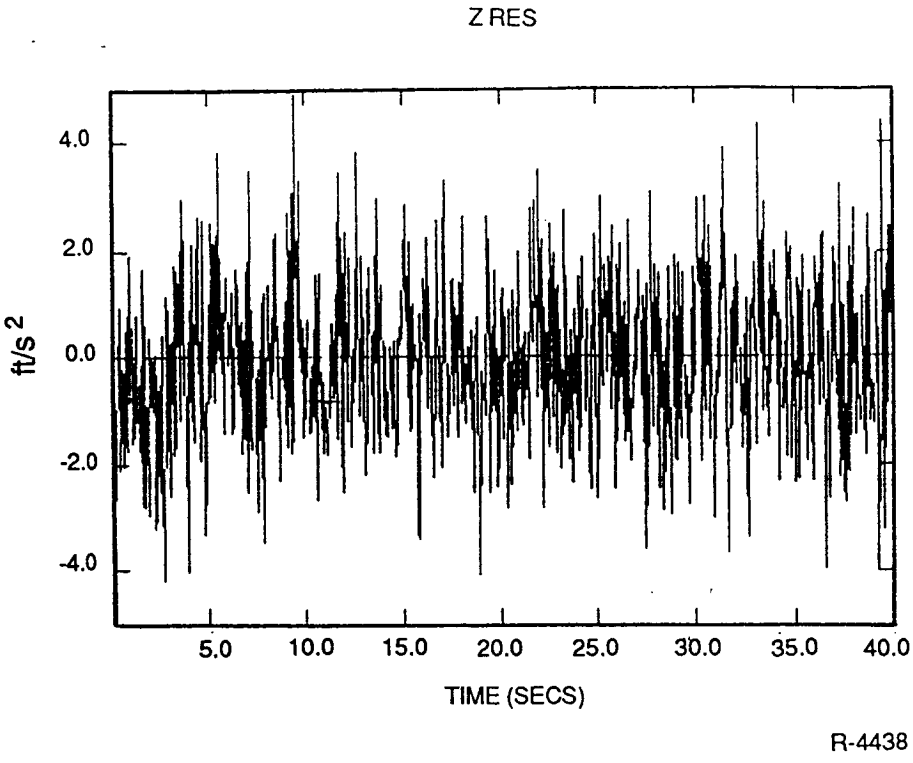


Figure 5-29. Z RES

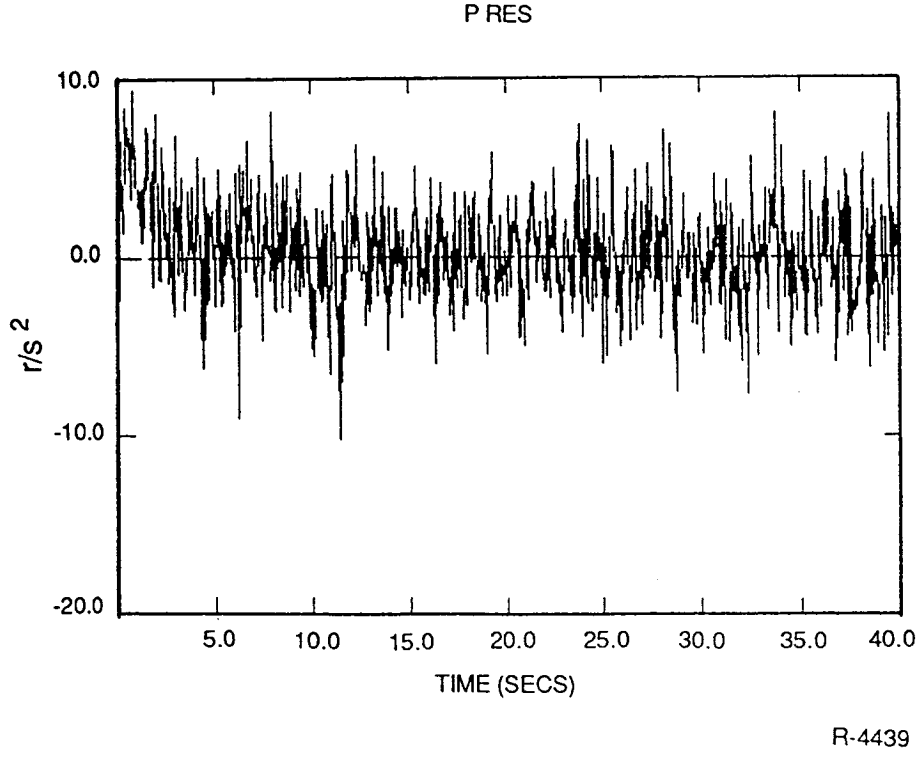


Figure 5-30. P RES

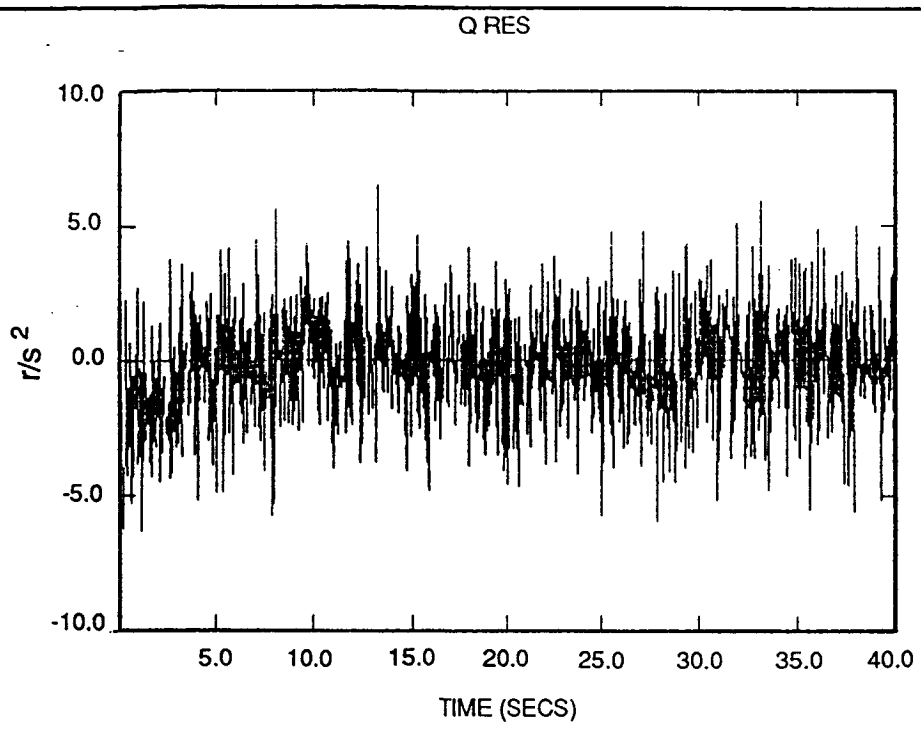


Figure 5-31. Q RES

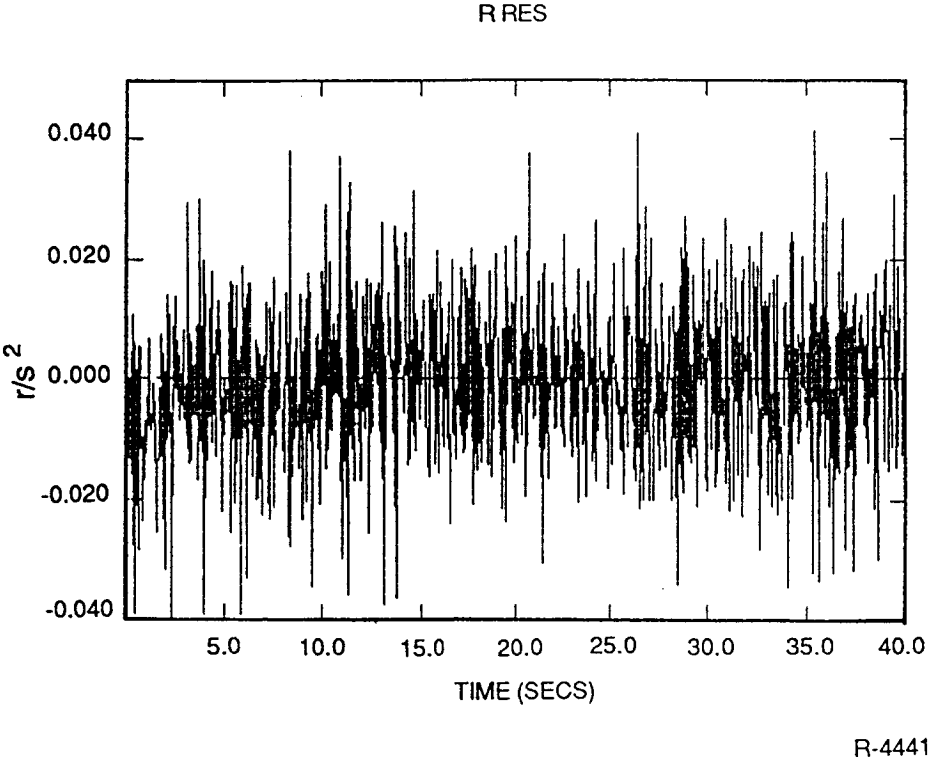


Figure 5-32. R RES

noise case. Figure 5-33 shows the trigger test for the left aileron.* Note that an initiation transient due to large low frequency errors causes the trigger to cross its threshold, however, an "inhibit" flag prevents decisions from being made for the first four seconds (two time constants of the HPFs) of operation. Other triggers look similar to Fig. 5-33. Low frequency errors which are excited during the maneuver are clearly evident, however, the design methodology's selection of thresholds account for the impact of these errors.

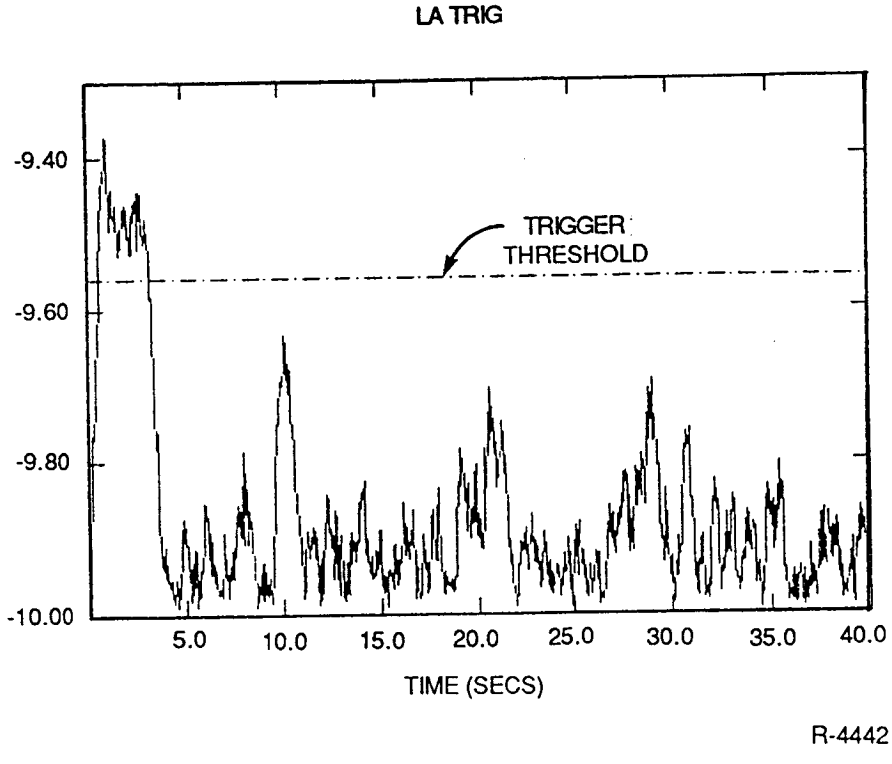


Figure 5-33. LA TRIG

*The scaling on all hypothesis test plots in this section do not correspond to the design data given in the previous subsection. This is because standard log-likelihood-ratio software was used to generate the simulation results. These log-likelihood-ratio tests are functionally equivalent to the ones specified above.

Climbing Turn Results (Detection of Totally-Missing Failures)- For these results, totally missing failures were similated (Effectiveness = 0 at 5.0 seconds before the initiation of the climbing turn maneuver at ten seconds). The simulations were run without turbulence but with sensor noise. No aircraft path engine failures were available. Both elevator and stabilizer failures were simulated, although no attempt was made to distinguish between the indistinguishable modes (same side elevator and stabilizer). Table 5-11 shows the trigger times and isolation times for each failure mode. Note that triggers for elevator and stabilizer failures occur before the maneuver because they are providing nonzero forces and moments to the aircraft during straight and level flight.

TABLE 5-11 TRIGGER AND ISOLATION TIMES FOR TOTALLY MISSING SURFACE FAILURES

Failure Mode	Trigger Time (sec)	Isolate Time (sec)
LS		6.15
RS		6.25
LA	10.9	11.70
RA	10.9	11.55
R		11.10
LE	5.4 (First)	Out of time (ambiguity group = LHT/RHT)
RE	5.4 (First)	Out of time (ambiguity group = LHT/RHT)

Figures 5-34 through 5-39 show the residuals for the rudder failure case. The trigger, verify, and isolate tests are shown in Figs. 5-40 to 5-47. Notice that the last isolation tests to decide are the hardest ones (as judged by Table 5-10), namely, R/RT and R/LT.

The elevator failure modes are not isolated (to the corresponding horizontal tail mode) during this run. This is due to the fact that the most difficult test (LHT/RHT) can not decide by its time limit. Figure 5-48 shows the LHT/RHT isolation test during both left and right elevator failures. Notice that triggers occur frequently and that in the greater majority of cases, the isolation statistics are heading in the correct direction. Examination of the elevator position indicates that its average value within the FDI bandwidth (.5 r/s to 6 r/s) is about six to seven degrees. Comparing this to the 7.4 degrees needed for distinguishability (Table 5-10; a factor of two times horizontal tail requirement due to 50 per cent effectiveness of elevator with respect to stabilizer), we see that the totally missing elevator failure for this maneuver is only marginally detectable (i.e., the likelihood of missing it is larger than desired). To see if, in fact, this failure mode is a critical failure, Fig. 5-49 shows the aircraft response during this failure. Comparison of Figs. 5-49 and 5-26 indicates that, in spite of the missing elevator, the climbing turn maneuver is successfully accomplished with little performance degradation. Thus, the qualitative assessment of this failure is that, for the control law being used, it is <u>not</u> severe.

Finally, Table 5-12 shows the trigger and isolate times for varying effectiveness levels of left aileron failures. The results are as expected. Smaller failures result in longer isolation times. When the failure is small enough, triggers occur, but isolation decisions can not be made (in this case

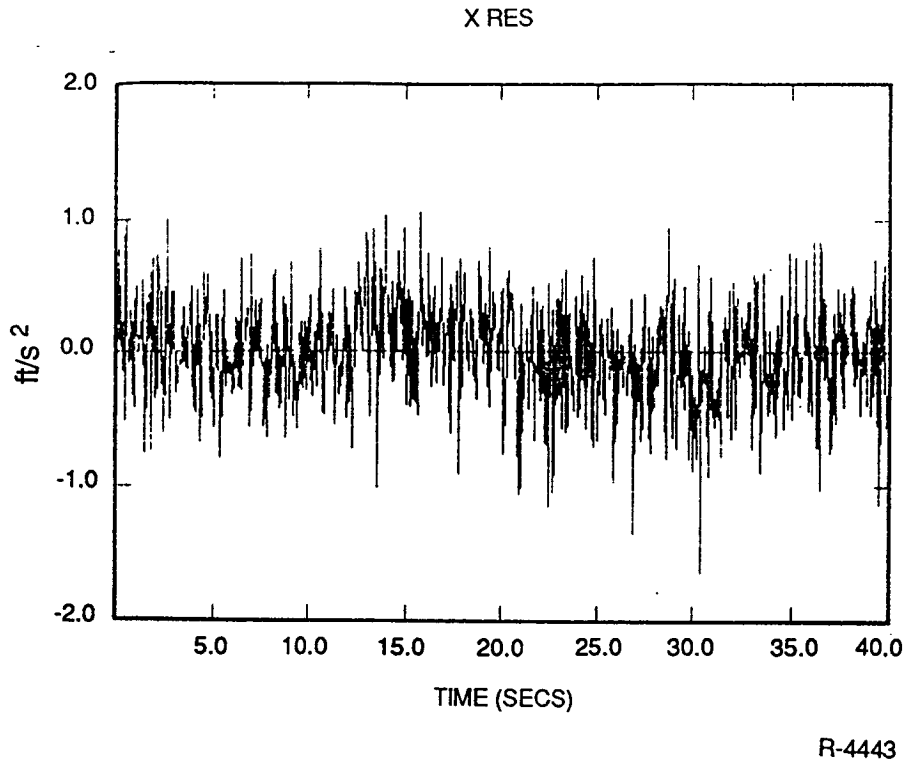


Figure 5-34. X RES

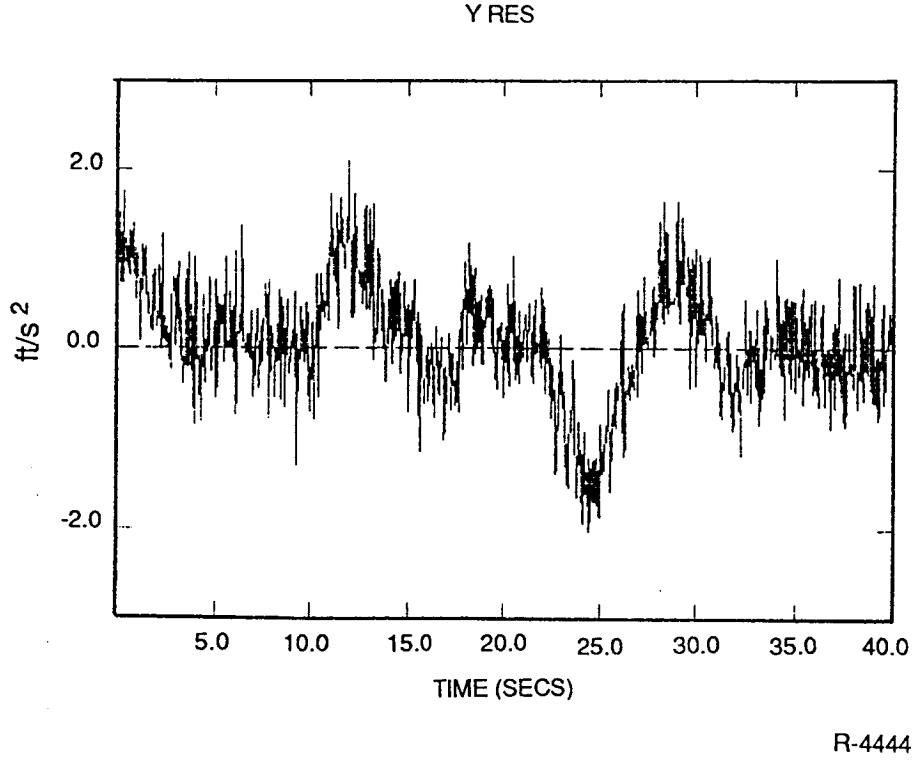


Figure 5-35. Y RES

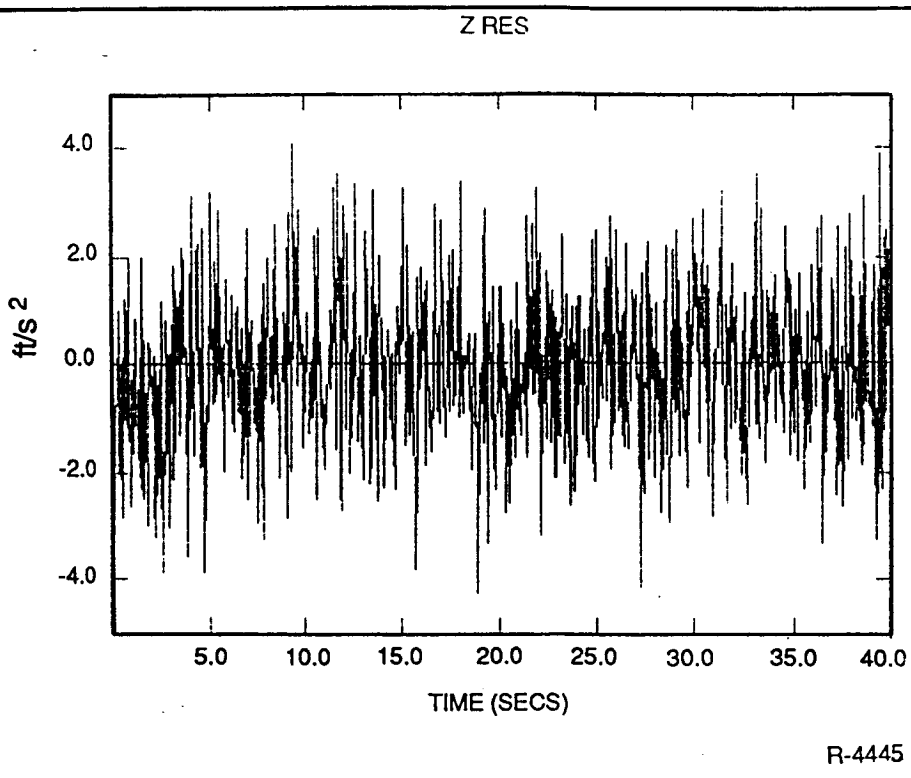


Figure 5-36. Z RES

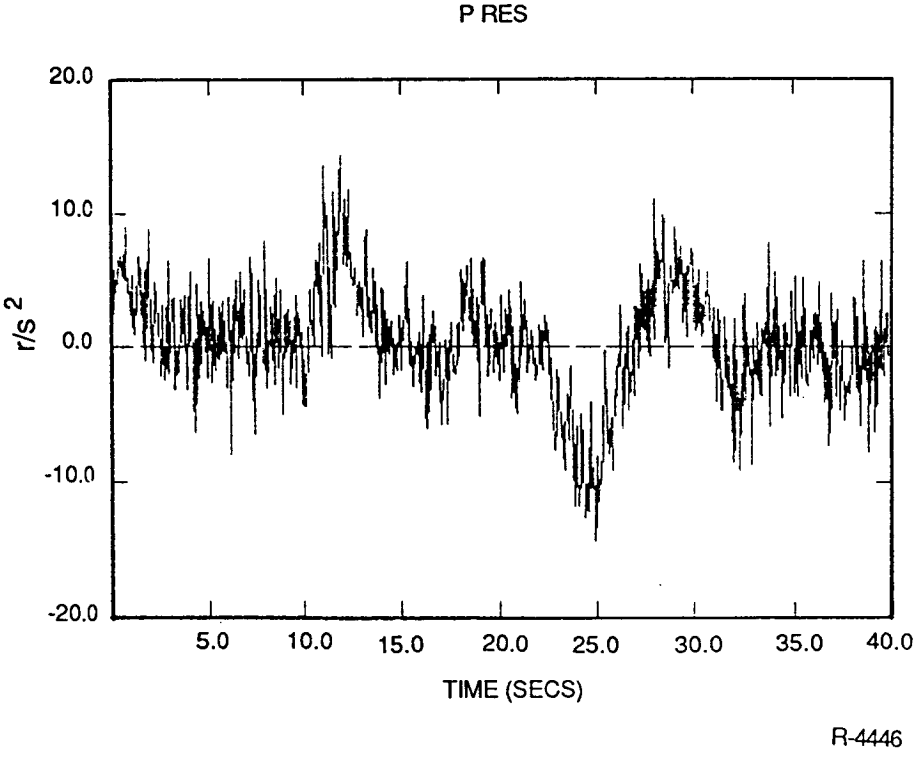


Figure 5-37. P RES

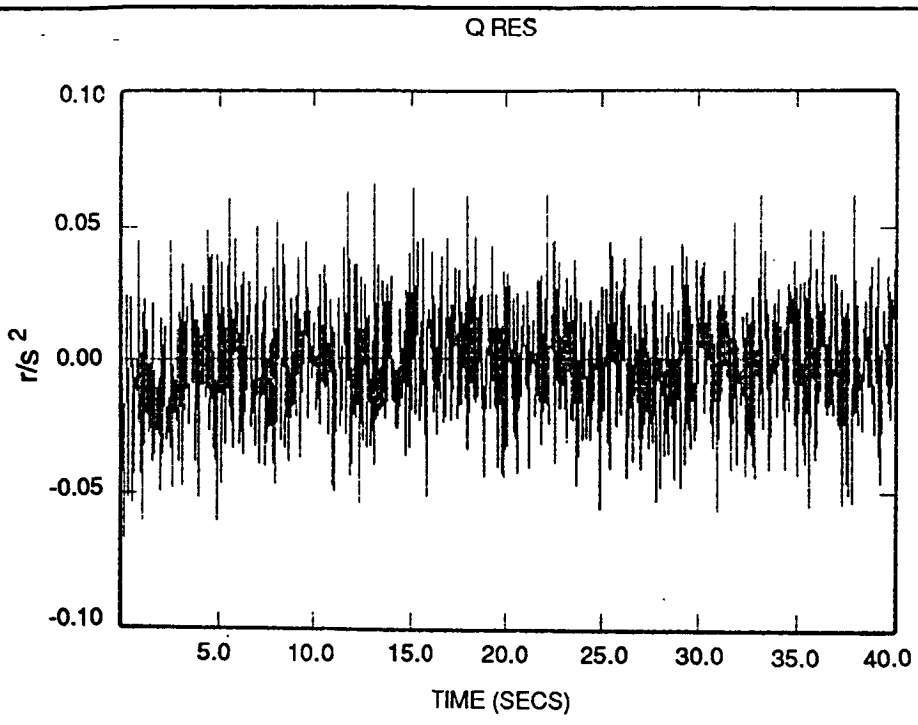


Figure 5-38. Q RES

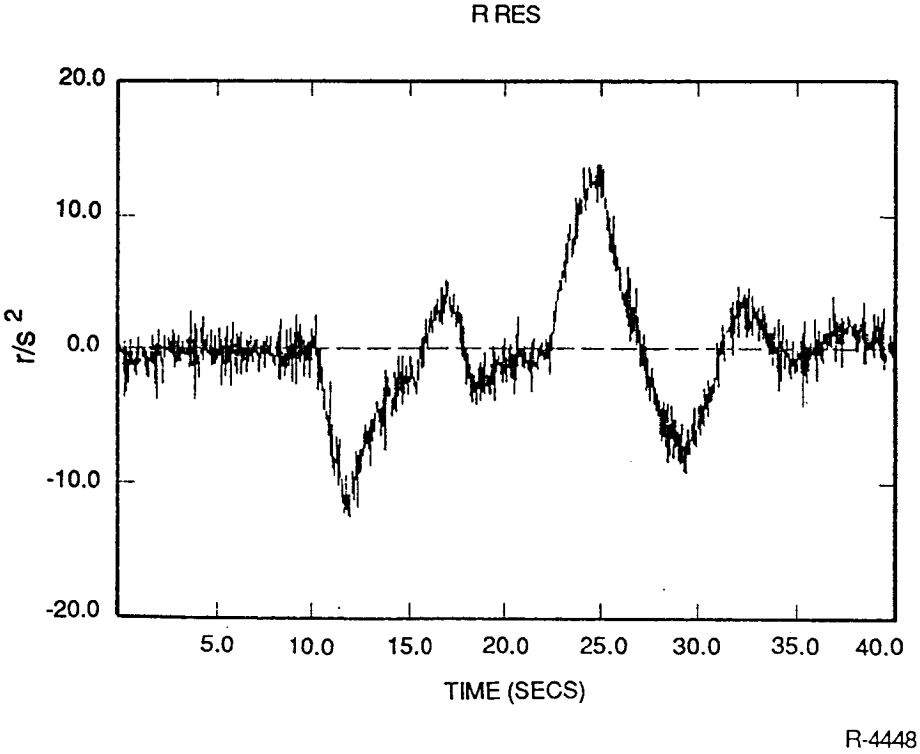


Figure 5-39. R RES

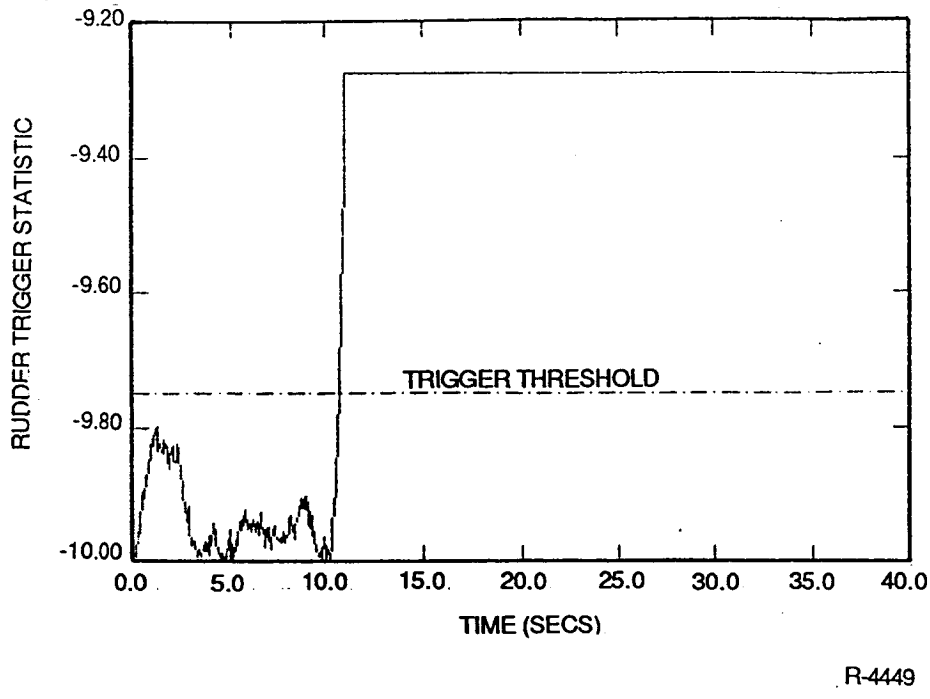


Figure 5-40. R TRIG

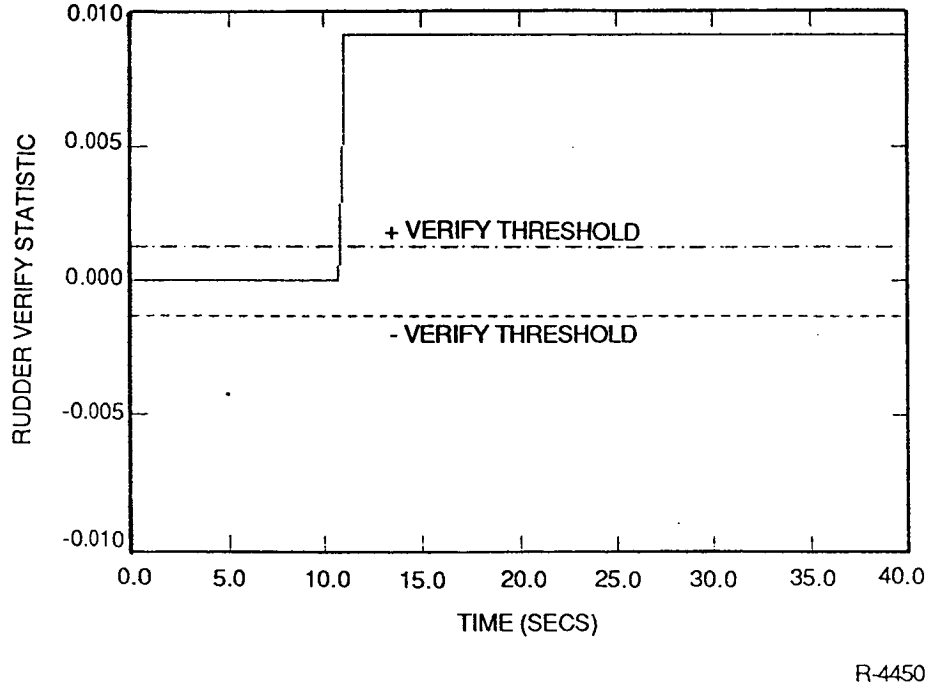


Figure 5-41. R VERF

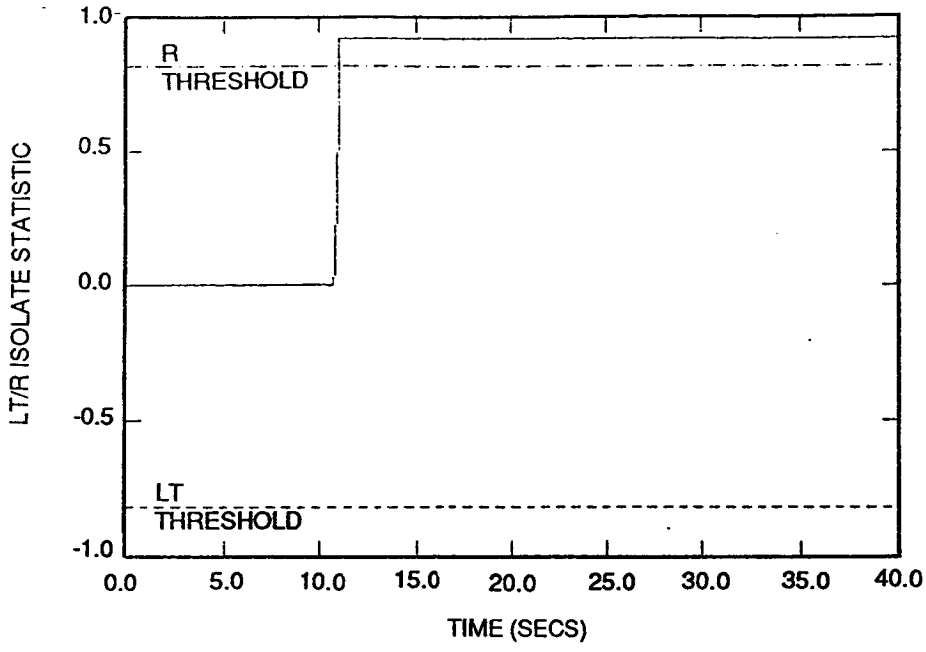


Figure 5-42. LT/R

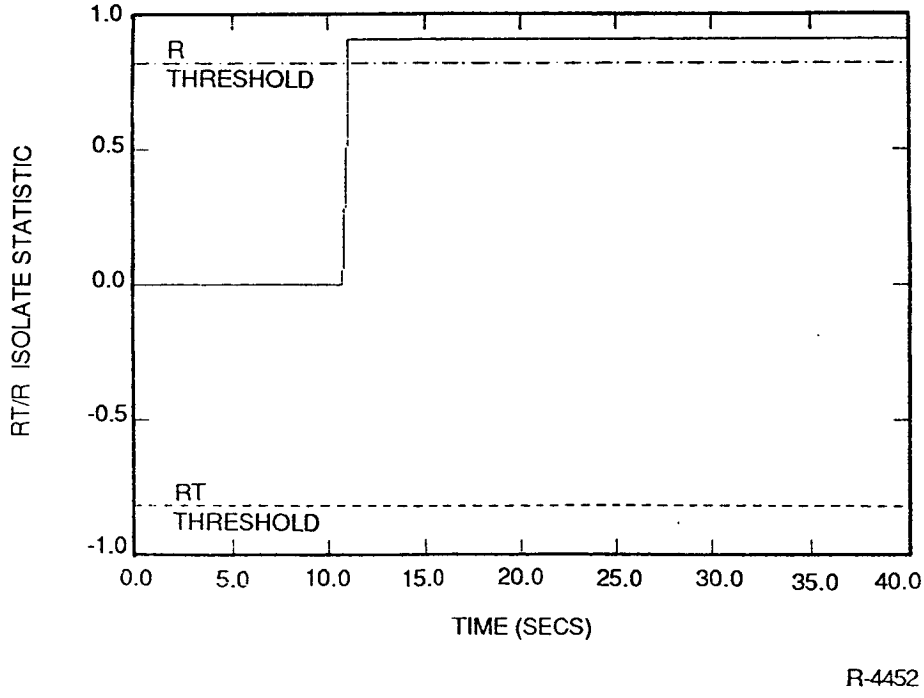


Figure 5-43. RT/R

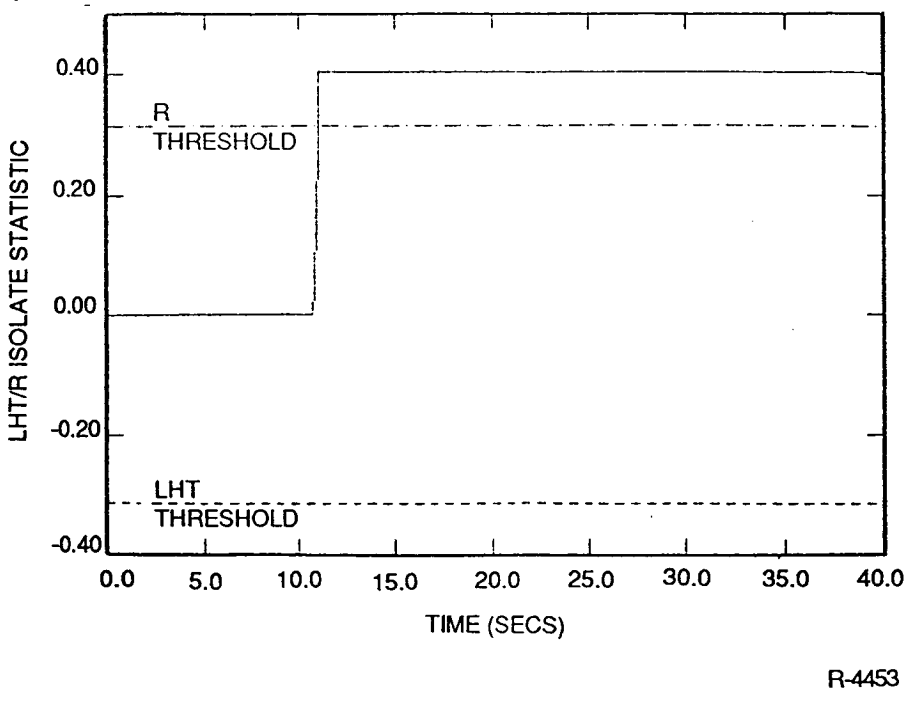


Figure 5-44. LHT/R

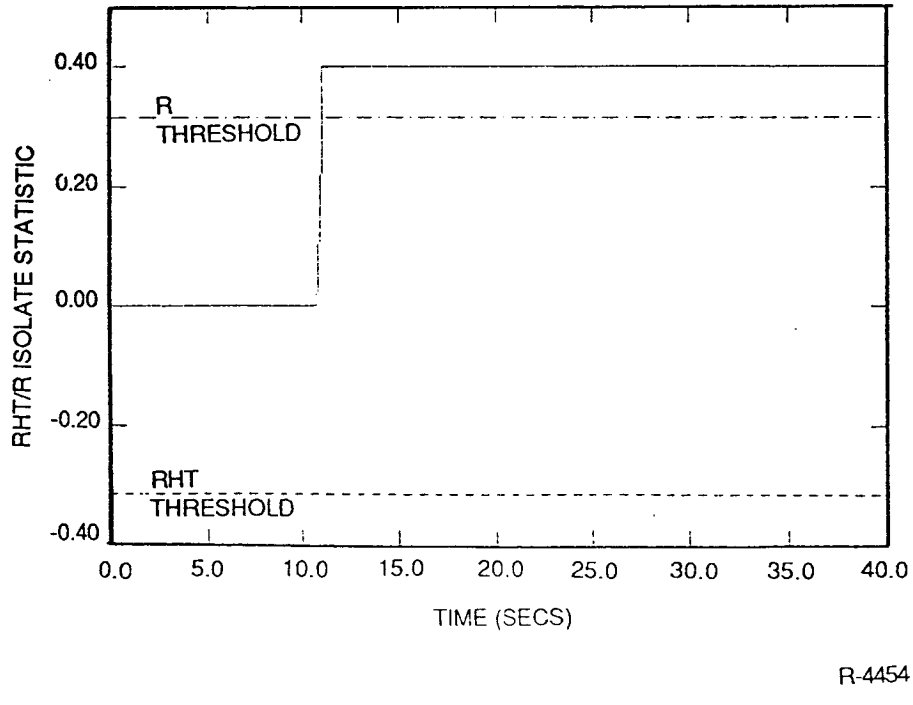


Figure 5-45. RHT/R

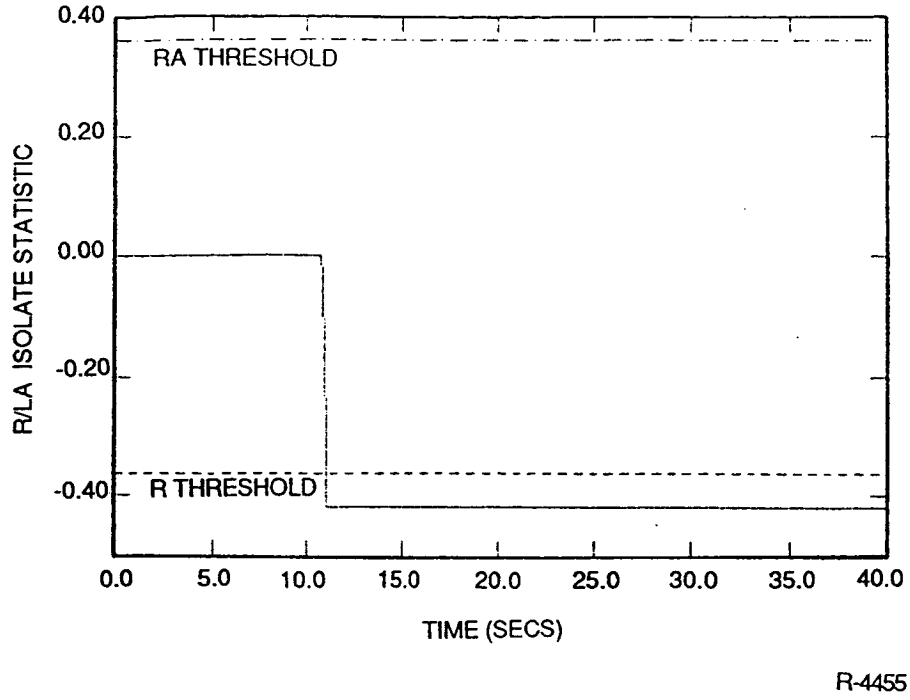


Figure 5-46. R/LA

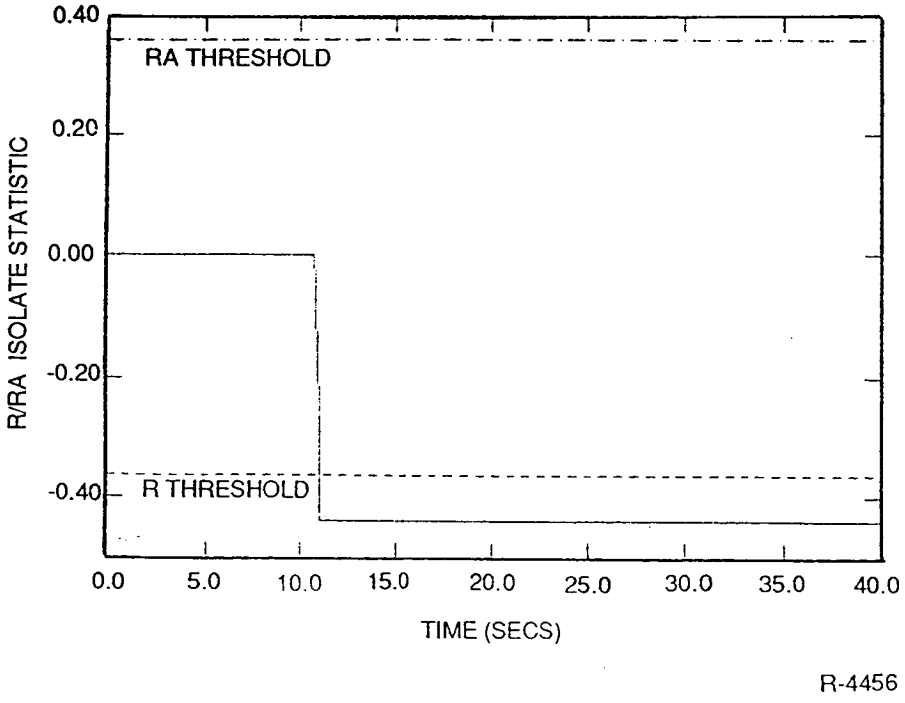
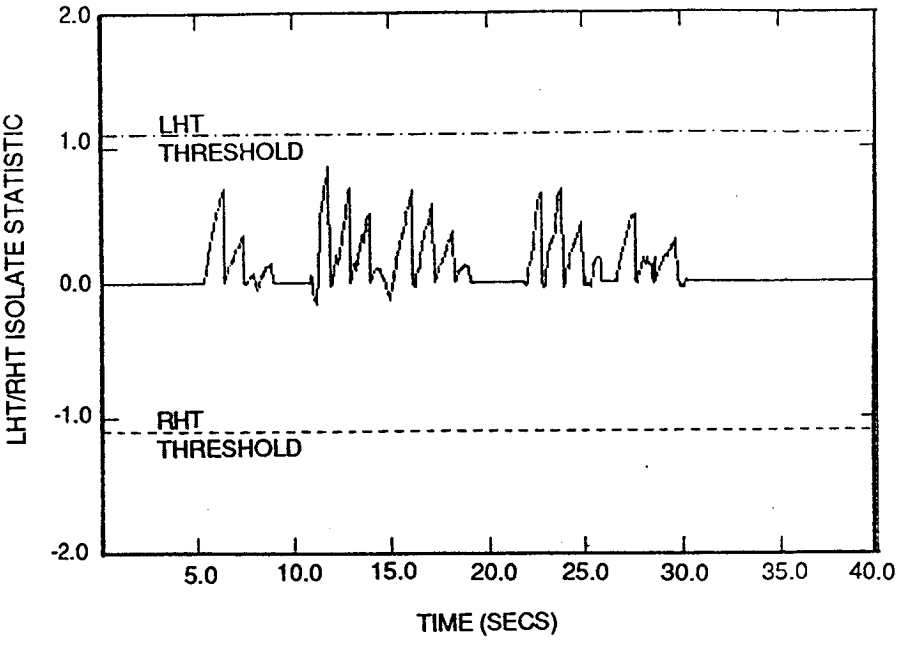
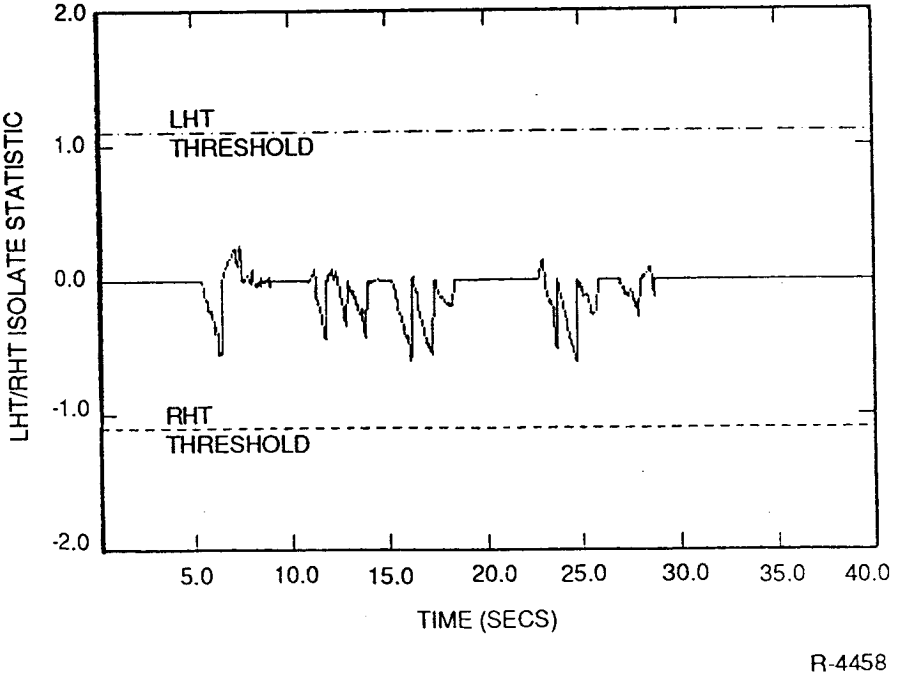


Figure 5-47. R/RA



(a). LHT/RHT Isolation Test, Left Elevator Failure



(b). LHT/RHT Isolation Test, Right Elevator Failure

Figure 5-48. 3/4

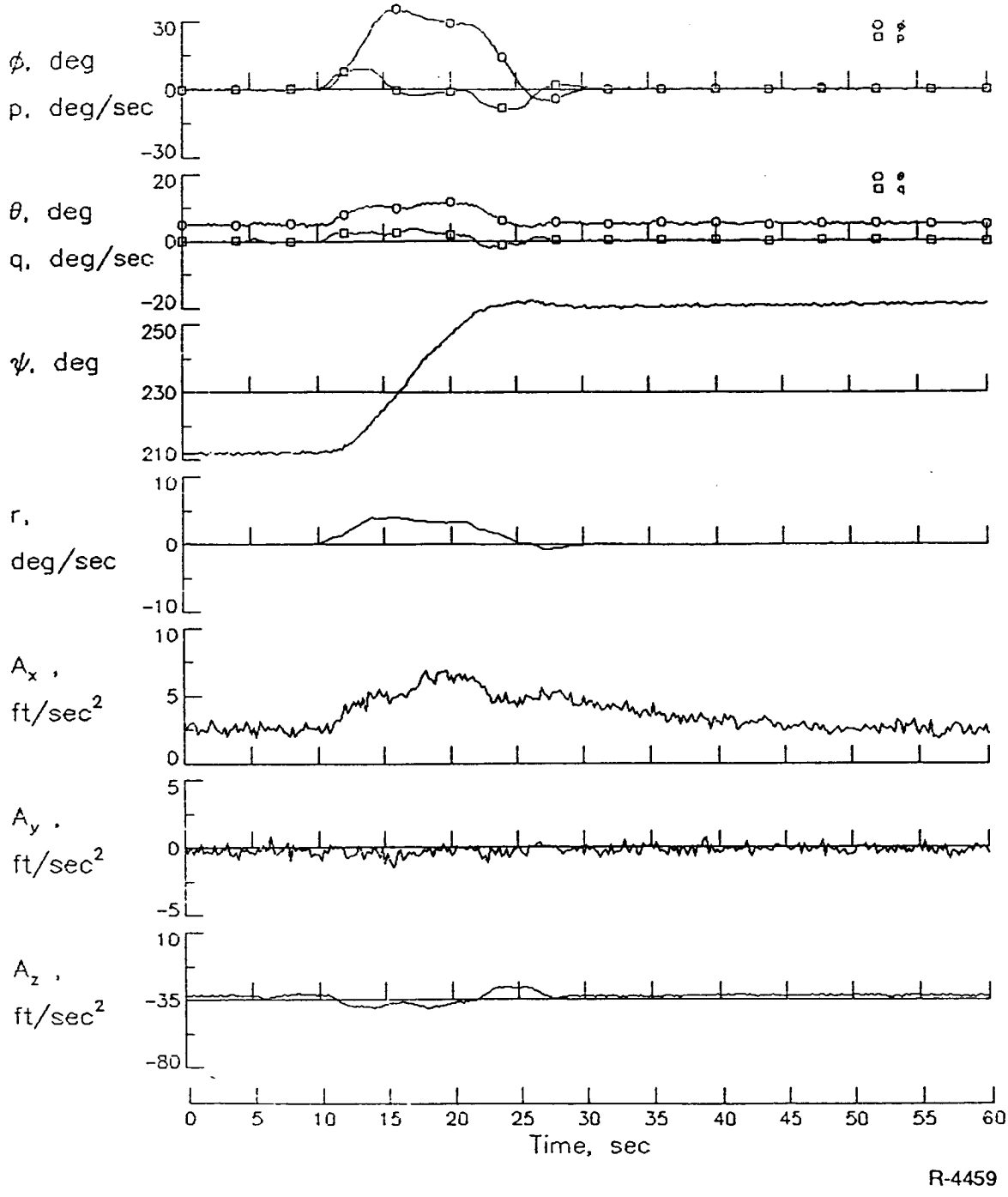


Figure 5-49. Maneuver Characteristics with Undetected Left Elevator Failure

due to the inability to verify the failure). When the failure is exceptionally small, no triggers are recorded.

TABLE 5-12. TRIGGER AND ISOLATE TIMES FOR LA FAILURES DURING CLIMBING TURN

LA Failure Size	Trigger Time (sec)	Isolate Time (sec)
100%	10.9	11.7
80	11.3	12.15
60	11.65 (First)	16.60
40	12.15 (First)	Verify Fails
20	None	None

Doublet Maneuvers - Several doublet maneuvers were simulated to test the algorithms reaction to severe maneuvers with a broad spectrum of excitation frequencies. The doublet maneuvers are defined by the following pitch and bank angle commands.

Pitch Doublet: At ~ 10 sec θ_c = 5 deg
At ~ 15 sec θ_c = -5 deg
At ~ 25 sec θ_c = 0 deg

Roll Doublet: At ~ 10 sec θ_c = 45 deg
At ~ 15 sec θ_c = -45 deg
At ~ 25 sec θ_c = 0 deg

These runs were made with no turbulence and sensor noise was added in a post-simulation processing function (i.e., not fed back to the commands in the control law). Figures 5-50 to 5-55 show the residuals for the roll doublet with no noise.

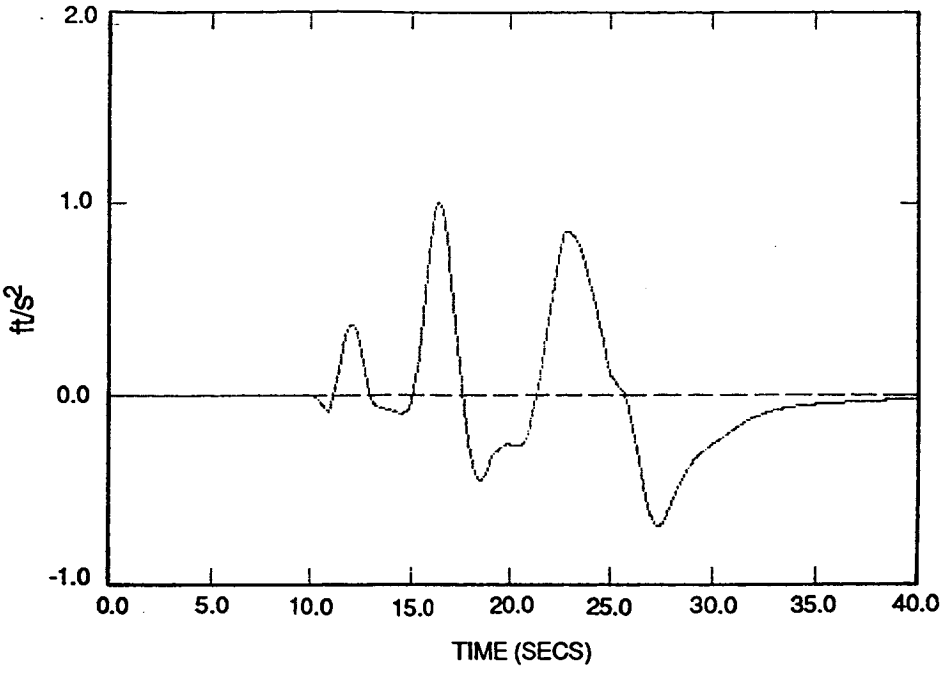


Figure 5-50. X RES

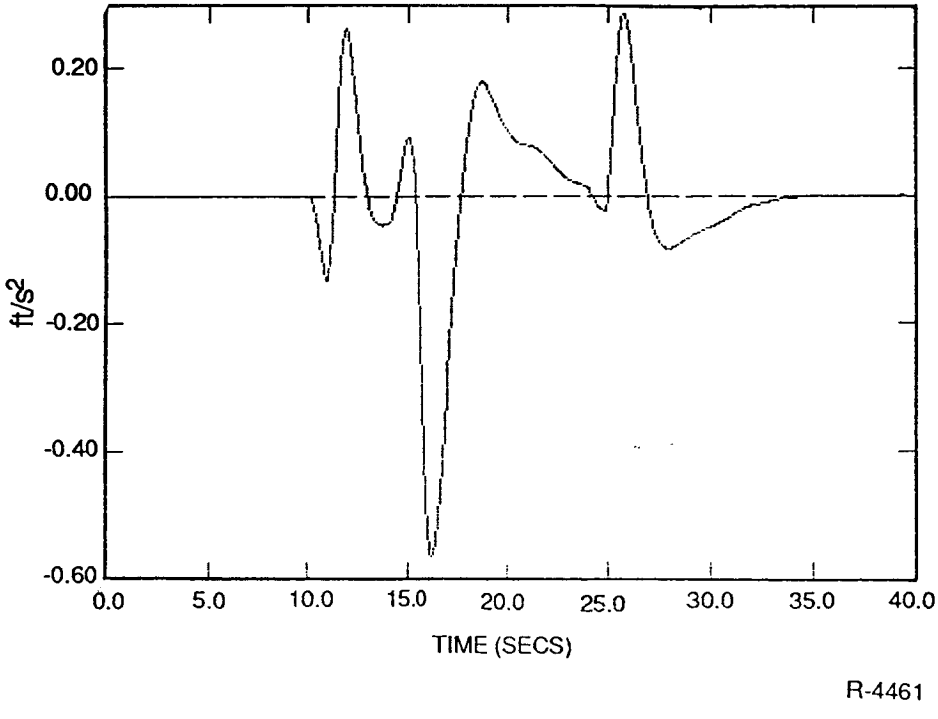


Figure 5-51. Y RES

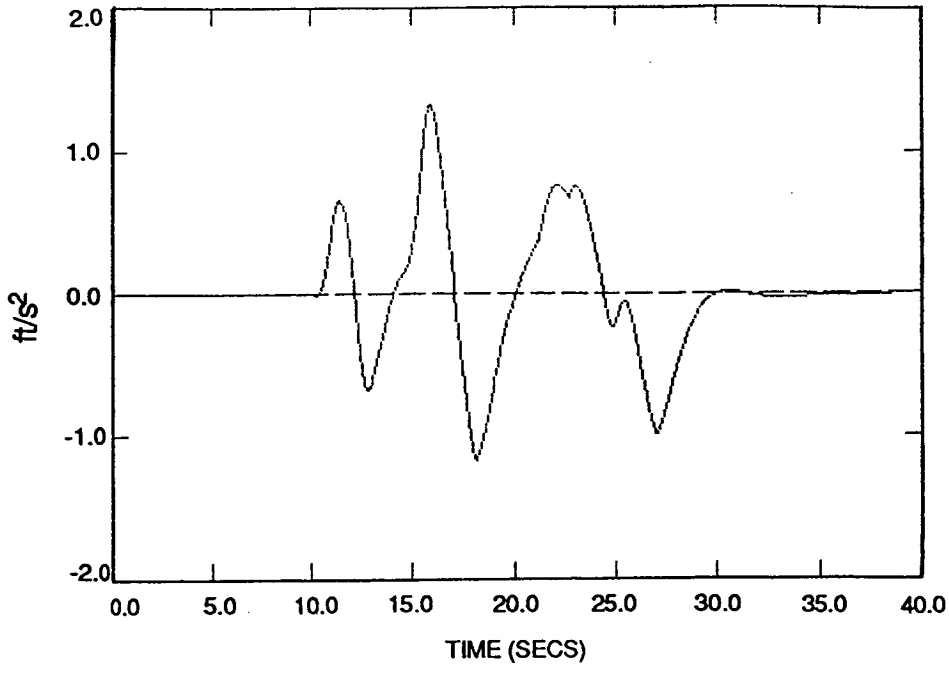


Figure 5-52. Z RES

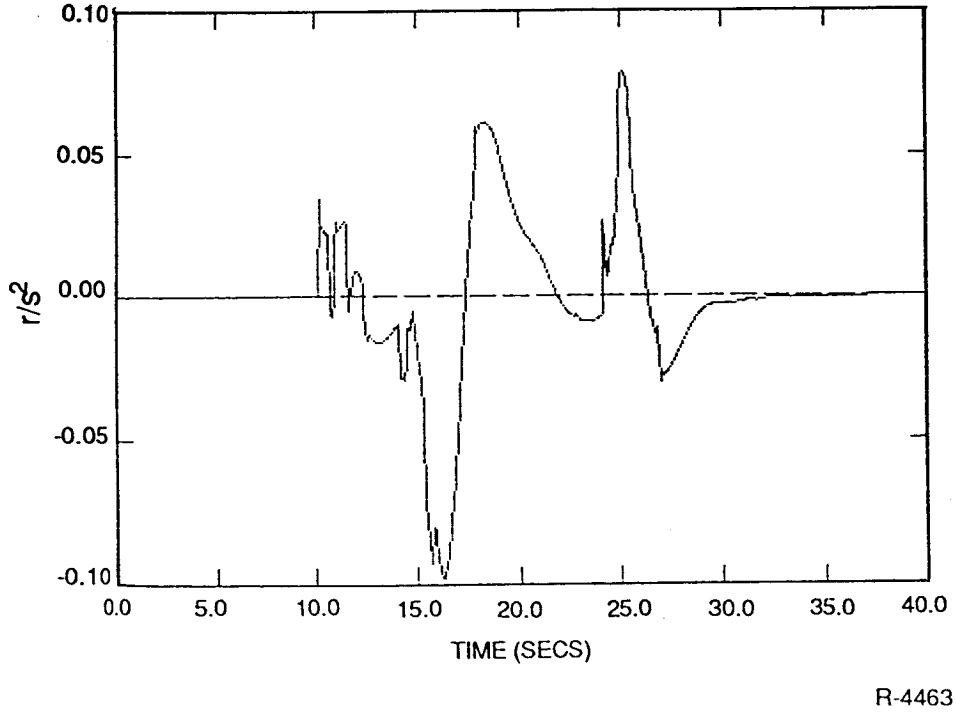


Figure 5-53. P RES

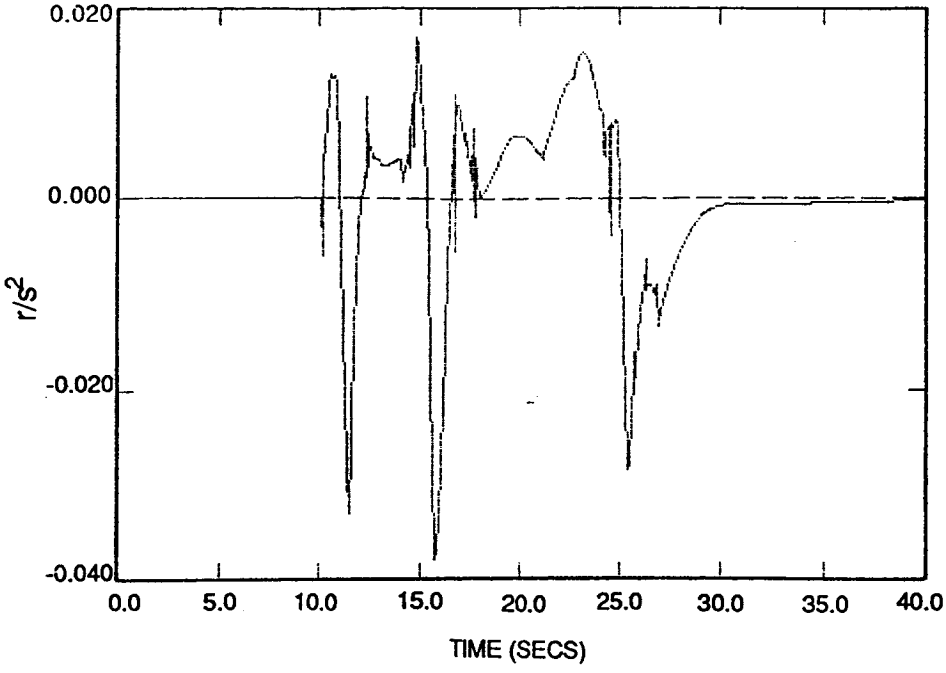


Figure 5-54. Q RES

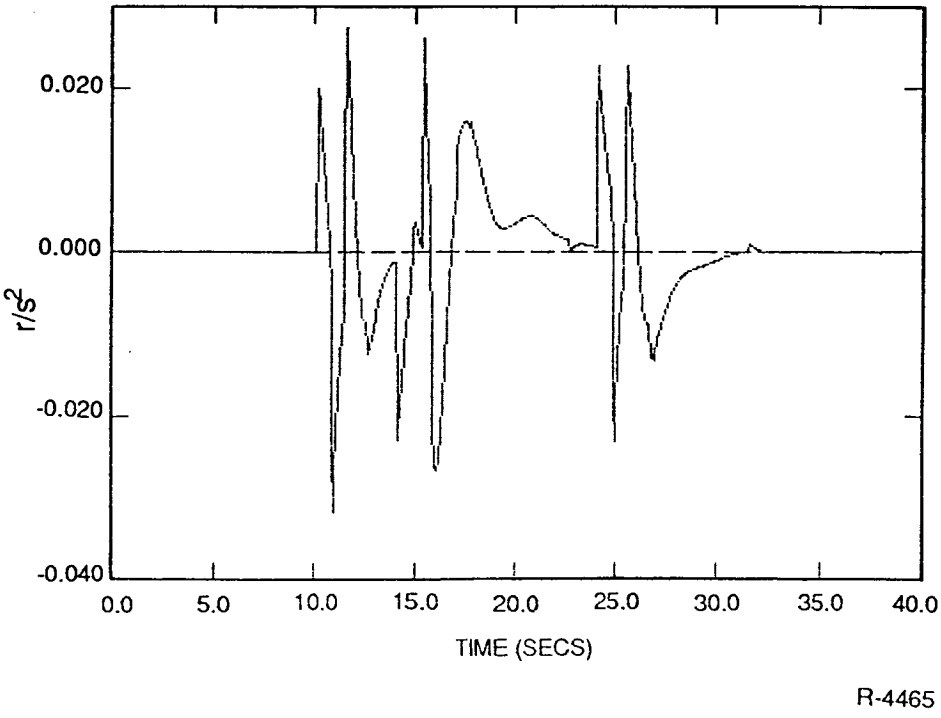


Figure 5-55. R RES

During the pitch maneuver with no failure, several LHT triggers occur, but the verify test either runs out of time or indicates a false trigger. Figure 5-56 shows the LHT trigger and verify tests for this case. For the roll doublet, triggers of every control element occur at one time or another, however, failures are not identified because the verify tests either run out of time or indicate a false trigger.

Examination of the residuals during these maneuvers indicates that the LF errors are substantially larger than those budgeted in the design. Thus, the results are not unexpected. In fact, it is surprising that in spite of the many false triggers, no failures are incorrectly declared. Apparently, the verify tests perform their function of declaring false triggers quite well during these maneuvers.

Forced Decisions - One of the options which was discussed in subsection 3.3 regarding the design of sequential tests was the possibility of making a fixed sample size decision when a sequential test's time limit is reached. It would be expected that smaller failures could be detected in this case, but that the likelihood of false alarms during severe maneuvers (e.g., the doublet maneuvers) would increase. Since 100 percent missing elevator failures could not be detected during the climbing turn maneuver, we decided to examine the results of this case when isolation decisions are "forced" at the time limits.

The forcing of decisions is accomplished by comparing the isolation statistic $S_I(N_I)$(see 5.1.3 for definitions) to zero. If it is larger than zero, the isolation test "passes" (decides in favor of H_i), and if it is less than zero, the test fails (decides in favor of H_j). Figure 5-48 indicates that

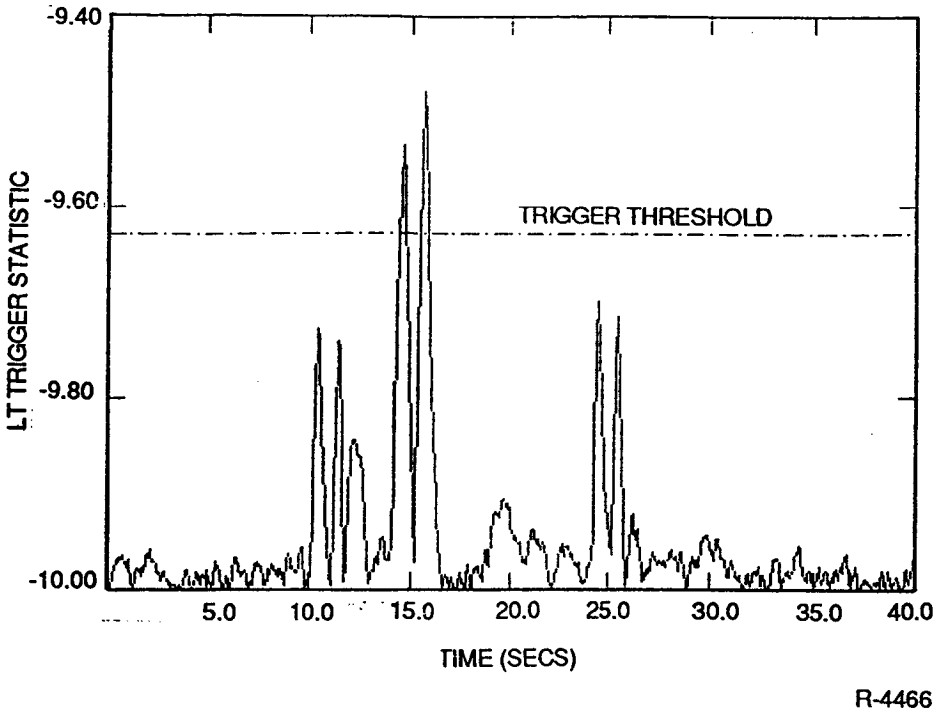


Figure 5-56 (a). LHT TRIG

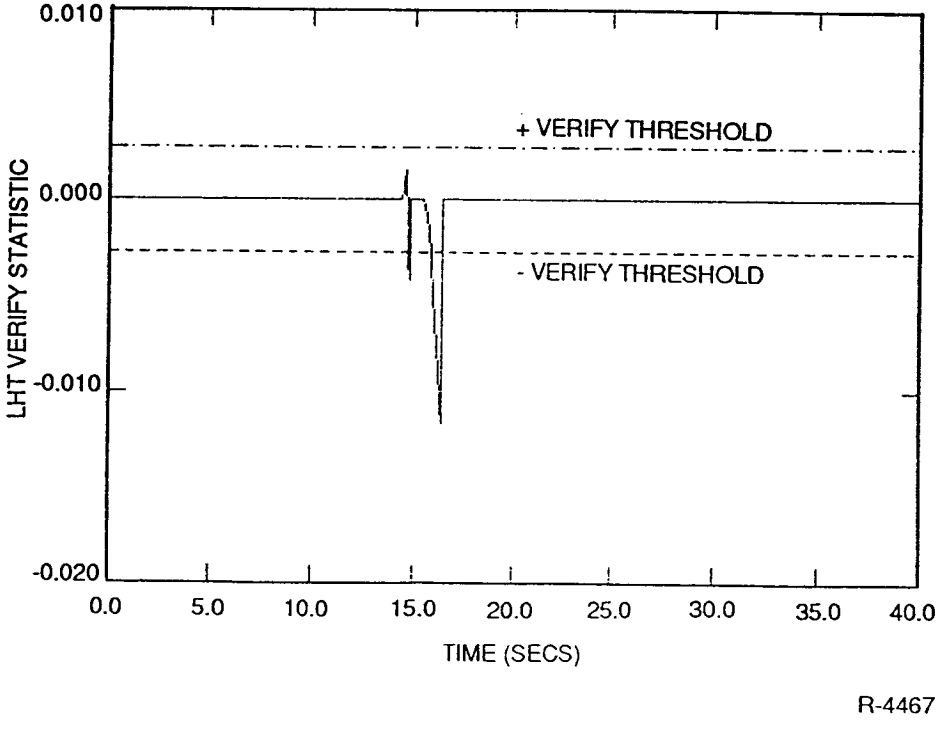


Figure 5-56 (b). LHT VERIFY

this procedure should work very well in the majority of trigger cases for the climbing turn maneuver. Simulation results for this case result in both left and right elevator failures being detected and isolated (to the corresponding LHT or RHT mode) by t=6.4 seconds. In the case of left elevator failures, the LHT/RHT test passes due to the forced decisions and for the right elevator failure, the LHT/RHT and RHT/R tests pass due to the forced decisions.

To see the impact of forced decisions on false alarm performance, the roll doublet was executed with no failures. Since many triggers occur, the likelihood of making an incorrect decision should be larger when isolate test decisions are forced at their time limits. However, no incorrect failure identifications are declared for this case. Part of the reason for this performance is that verify test decisions are not forced. Thus, many false triggers are indicated due to verify tests being unable to decide at their time limits. If verify decisions were also forced, it is possible that false alarms could have been declared.

Threshold Scheduling - Methods for scheduling both fixed sample size tests and sequential tests were discussed in subsection 3.3. These methods were applied to the aircraft-path FDI system to see if false triggers during severe maneuvers could be avoided.

Basically, threshold scheduling is accomplished through the use of failure insensitive estimates of upper bounds on the size of the residuals due to errors which were not accounted in the nominal design. In the beginning of this subsection, we noted that improved "nominal" performance (in terms of detectability levels) could be achieved by reducing the design envelope and scheduling thresholds so that periods of potentially large error do not cause

incorrect decisions. Thus, a natural choice of scheduling parameters would be of the form,

$$E_i(k) = \sum_j | K_{ij} y_j | \qquad (5\text{-}71)$$

where j is summed over the measurements which were ignored in the error budget and K_{ij} represents an upper bound on the error in the corresponding dimensional coefficient. (Note that since these errors represent errors in the residuals, and the residuals are high pass filtered, the "error terms" must also be high pass filtered). Unfortunately, the results obtained with such a procedure are not adequate. Figure 5-57 shows the LHT trigger test for the pitch doublet maneuver with no noise or turbulence. The threshold schedules correspond to five percent errors in the coefficients which were ignored in the error budget. Clearly Eq. 5-71 does not characterize the errors which are being experienced during this maneuver.

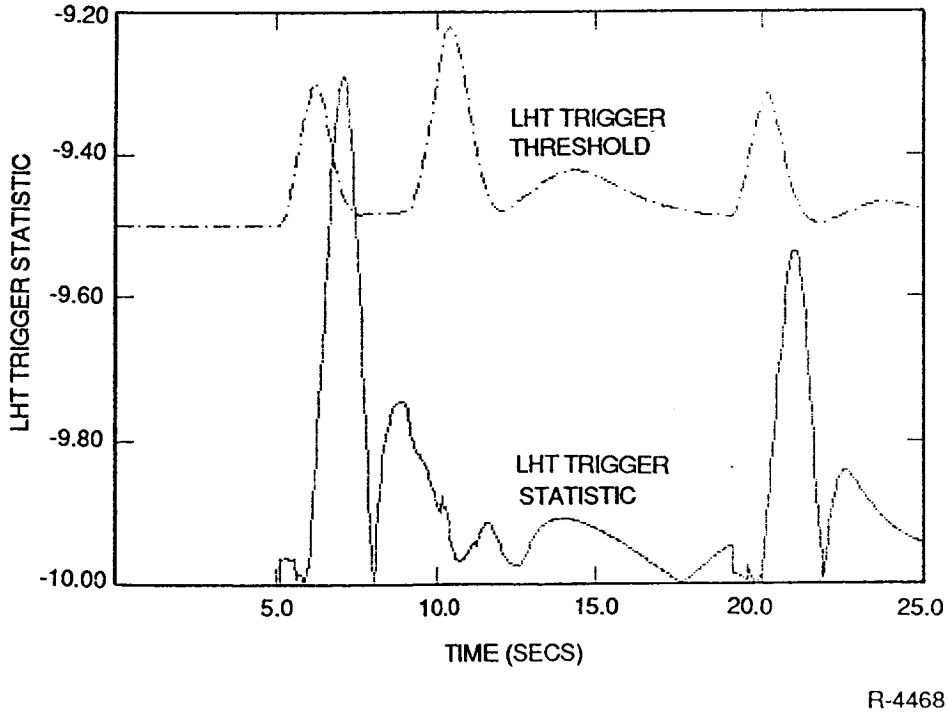


Figure 5-57. LHT Trigger Threshold and Statistic Threshold Scheduled as a Function of "Ignored Errors"

Several attempts were made to find the proper signals upon which threshold scheduling could be based. These included the use of deflection measurements, acceleration measurements, angular rates, and others. No combination was found which satisfied the requirement of eliminating false triggers during all doublet maneuvers. More investigation of these issues is, therefore, necessary. However, we note that in the search for appropriate threshold schedules, it was observed that errors in the longitudinal residuals may be correlated with $\dot{\alpha}$. Since no $\dot{\alpha}$ term was modeled in the residual generation equations, it is possible that this is an important source of error.

Discussion - There are several conclusions about the aircraft-path FDI algorithm and its design methodology which can be drawn from these tests. Broader conclusions about the entire effort are given in Section 6. The conclusions drawn here are based on the results which have been presented as well as a more detailed analysis of several simulation results.

In general, we feel that the aircraft path FDI algorithm performed quite well, and as expected, under conditions for which it was designed. It performed surprisingly well during conditions in which errors were substantially larger than those included in the design procedure's error budget.

Room for improvement, of course, still exists, and several comments in this regard are given below.

1. Although no engine failure tests were run, examination of throttle commands indicates that some engine failure modes (e.g., stuck: in the aircraft path) may produce lower frequency signatures than the other failure modes. In order to accommodate such modes, it may be desirable to reduce the bandwidth of the HPF only for tests involving the engine failures.

2. Increased sensitivity to failures and greater flexibility in the design of isolation tests may be achievable if the full matrix of

isolation tests (i.e., "pure" rejection tests) is implemented. This is as opposed to the combined test described by Fig. 5-25.

3. Decreased decision delays may be achievable by starting the sequential tests prior to the trigger time. This process would require storage and processing of residuals over the "rollback" interval. The disadvantage of such an approach is that the same errors which might create false triggers could then be present in the verify tests (i.e., the verify test is not an independent test).

4. In keeping with the desire to make the verify test independent to reduce false alarms, we could reinitialize the HPFs at the trigger time. Although this would remove the effect of large LF errors which cause false triggers from the verify test, it would also remove the effect of failures. Continued excitation during the sequential test interval is then required. This procedure might be most effective when a "rollback" interval is used.

5. The FDI interval and the length of time in which the failure signatures remain constant are close enough to warrant investigation of noncoherent processing schemes. That is, trigger, verify, and isolate tests might perform better if they were based on sums of projected and squared residuals.

6. The sequential test design procedure seems to produce a relatively conservative choice of thresholds when errors are within the budget. This conclusion is based on observing that, in cases where failures are not correctly detected, the isolation tests which run out of time are clearly heading in the direction of a correct decision. Some further refinement and/or analysis of the current design procedure is desirable.

7. When high frequency unmodeled errors are present, it is possible that the likelihood of making incorrect decisions early in a sequential test could increase. If adequate high frequency error models are used in the design procedure, the thresholds will increase to accommodate this. However, an alternative scheme which might allow greater sensitivity in the sequential tests would be to start the sequential test with large thresholds and reduce them as the high frequency errors become less important. No formal theory or heuristic design procedure exists for such a concept and if important, the concept should be investigated further.

8. More work is needed to define meaningful threshold schedules for operation of the algorithm during severe maneuvers.

9. The moving window average nature of the trigger tests is a very inefficient implementation of a low pass filter (LPF). It was chosen on the basis of the optimal fixed sample size hypothesis tests. Other LPF filter structures which achieve similar noise rejection and perhaps have a better transient response with less computation may be of interest.

5.2 ACTUATOR PATH FDI ALGORITHM DEVELOPMENT

Two distinct actuator path FDI algorithms were investigated for this project. The first algorithm is a "fixed" or constant threshold algorithm which maintains a desired false alarm rate during worst case model error at the expense of reduced failure sensitivity during all phases of operation. The second algorithm uses the threshold scheduling concepts developed in subsection 3.3 to achieve reduced failure sensitivity only during times when large error is expected. This latter approach was demonstrated on only two surfaces (left and right aileron) because errors other than those assumed in this approach were evident in the other surfaces. The development of both algorithms and corresponding simulation results are described below.

5.2.1 Residual Generation

Both actuator FDI algorithms described in this section are based on the same residual generation mechanism. For this project, "open-loop" (see subsection 3.2) actuator residuals are generated by passing the commanded control value through an actuator model to produce an estimate of the "actuator" output position and subtracting this estimate from a measurement of the actuator output. Although other residual generation mechanisms are possible (e.g. finite memory or closed loop residuals) it is felt that this mechanism provides the best ratio between failure sensitivity and model error effects.

The actuator model (and, hence, what we mean by an actuator) is defined by the locations of actuator commands and output measurements. Table 5-13 shows the definitions used for this project. These definitions assume that the only input which drives each control element is computed by a digital flight control system (DFCS) that incorporates both stability augmentation and

TABLE 5-13. B-737 ACTUATOR DEFINITIONS

SURFACE	COMMAND	OUTPUT	UNITS
L&R Stabilizer	DFCS Command	Position at Hinge	degrees
L&R Elevator	DFCS Command	Position at Hinge	degrees
Rudder	DFCS Command	Position at Hinge	degrees
L&R Aileron	DFCS Command	Position at Hinge	degrees
L&R Throttle	DFCS Command	EPR (Scaled)	k-lbs

pilot inputs. Thus any failure between the DFCS output and the control measurements is considered as an actuator failure. Control surface deflection measurements are taken on the hinge (possibly using an LVDT on a control rod or an RVDT on the hinge itself).

The estimated control value (for the "next pass") is obtained using a generic actuator model which incorporates linear first order dynamics, rate and position limits, and static cable stretch. The sequence of operations used to generate the control estimate and residual for each actuator are,

Dynamics:
$$\hat{\delta}_k = e^{-\omega\Delta}\hat{\delta}_{k-1} + (1-e^{-\omega\Delta})\ \delta_c(k-1) \tag{5-72a}$$

Rate Limiter:
$$\text{IF } [\hat{\delta}_k - \hat{\delta}_{k-1}]/\Delta > \text{RLIM}$$

$$\text{THEN } \hat{\delta}_k = \hat{\delta}_{k-1} + \text{sgn}(\hat{\delta}_k - \hat{\delta}_{k-1}) \cdot \text{RLIM}\cdot\Delta \tag{5-72b}$$

Cable Stretch:
$$\hat{\delta}_k = \frac{1}{1 + \bar{Q}\cdot SF}\ \hat{\delta}_k \tag{5-72c}$$

Position Limit:
$$\text{IF } (\hat{\delta}_k > \text{Max}) \quad \text{THEN } \hat{\delta}_k = \text{Max}$$

$$\text{IF } (\hat{\delta}_k < \text{Min}) \quad \text{THEN } \hat{\delta}_k = \text{Min} \tag{5-72d}$$

Residual:
$$\nu_k = \delta_m(k) - \hat{\delta}_k \tag{5-72e}$$

where δ_m denotes the control measurement, $\hat{\delta}$ denotes the estimate, Min and Max are the position limits, $\bar{Q}$ is dynamic pressure, SF is a stretch factor, RLIM is the rate limit, ω is the first order bandwidth and sgn is the signum function. The parameters of this model for each control element are given in Table 5-14. The sample time, Δ, is .05 seconds. The elevator command and its corresponding measurement is assumed to be given relative to the mechanical trim while all other surface positions are relative to some body oriented coordinate frame. For the throttle, different bandwidths are used for spool up ($\delta_c > \hat{\delta}_{k-1}$) and spool down ($\delta_c < \hat{\delta}_{k-1}$).

TABLE 5-14. ACTUATOR MODEL PARAMETERS

SURFACE	ω	RLIM	SF	MIN, MAX	UNITS
L&R Stabilizer	1.5 r/s	±10 deg/s	0.0	-14, 3	degrees
L&R Elevator	22 r/s	±20 deg/s	.0023	-10, 10	degrees
Rudder	22 r/s	±18 deg/s	0.0	-10.3, 10.3	degrees
L&R Aileron	20 r/s	±20 deg/s	.0016	-20, 20	degrees
L&R Throttle	2.0 r/s (spool up) 1.0 r/s (spool down)	None	0.0	(10), (60)	k-lbs

Several error sources which result in non-zero residuals are present. The first, and most obvious is sensor errors. These include bias and sensor noise as detailed in Table 5-15. Scale factor errors result from errors in SF. Errors which are excited by high frequency commands are present due to inaccuracies in the dynamic model. Finally, rate limit errors may be present. This last error is particularly true for the rudder in which the actual rate limit is a nonlinear function of surface position. Also, the throttle command

TABLE 5-15. SENSOR ERRORS

SURFACE	NOISE (1-σ)	BIAS (MAX)	UNITS
L&R Stabilizer	.1	.1	degrees
L&R Elevator	.1	.1	degrees
Rudder	.1	.1	degrees
L&R Aileron	.1	.1	degrees
L&R Throttle	(.01)	(.02)	k-lbs

to EPR model may have scale factor errors due to uncertainty in the scaling of both commands and EPR to thrust units (klbs).

Simulation runs using the maneuvers defined in subsection 5.1.4 were made. Figures 5-58 to 5-66 show the residuals for each control element during the roll doublet maneuver with no noise or failures. In these figures, the initial bias is removed. There are clearly many error sources besides sensor noise present. For the ailerons and elevators errors occur mainly during the command steps which initiate each phase of the maneuver. These errors are consistent with dynamic inaccuracies at high frequencies. The stabilizer errors have a lower frequency content than those of the elevators and ailerons. These errors are also consistent with high frequency dynamic errors since the bandwidth of the stabilizer is substantially lower. For the errors observed in the rudder and throttle, it is likely that scale factor errors and the existence of various nonlinearities in the simulated engine models play a large role in contributing to residual errors. For the rudder, it was discovered that nonlinear rate limits are used in the simulation to account for varying aerodynamic loads during surface motion. The error between this nonlinear rate limit and the constant rate limit used in Eq. 5-72b is evident in Fig. 5-62.

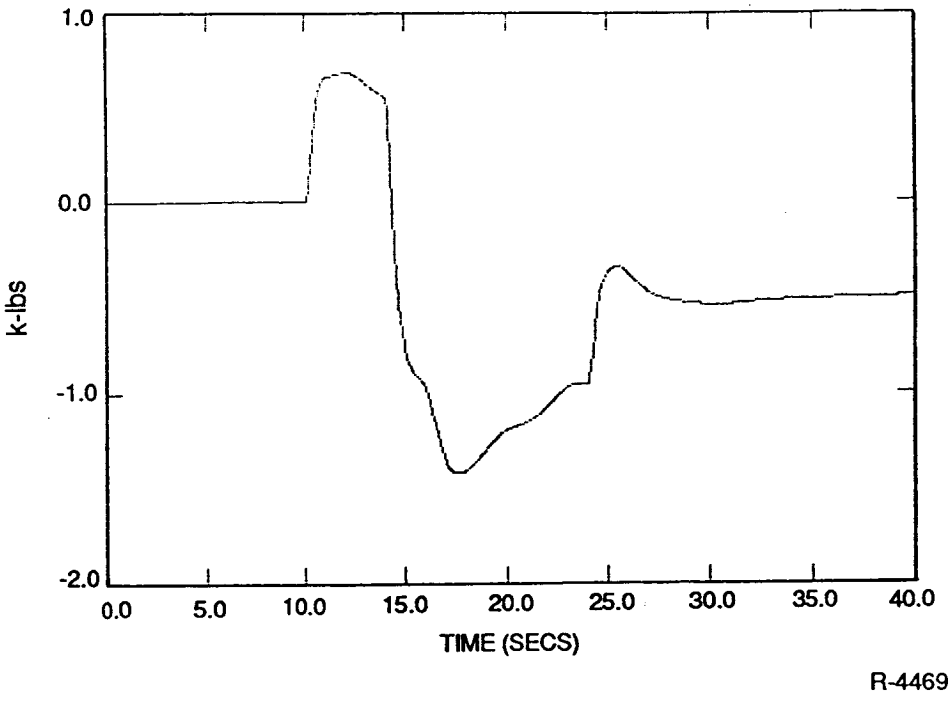


Figure 5-58. LT RES

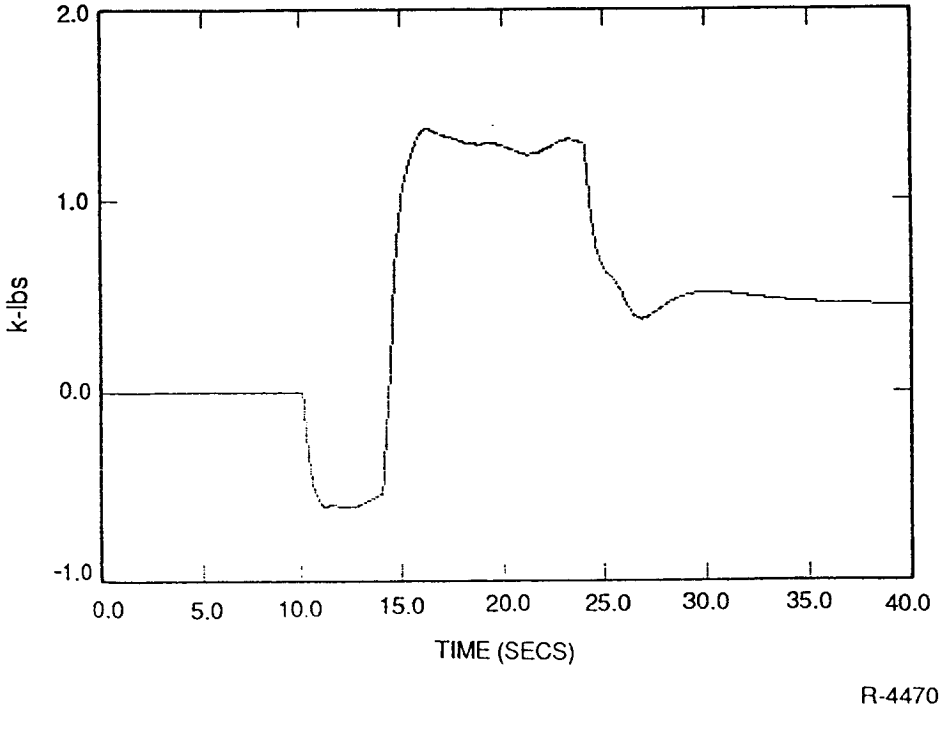


Figure 5-59. RT RES

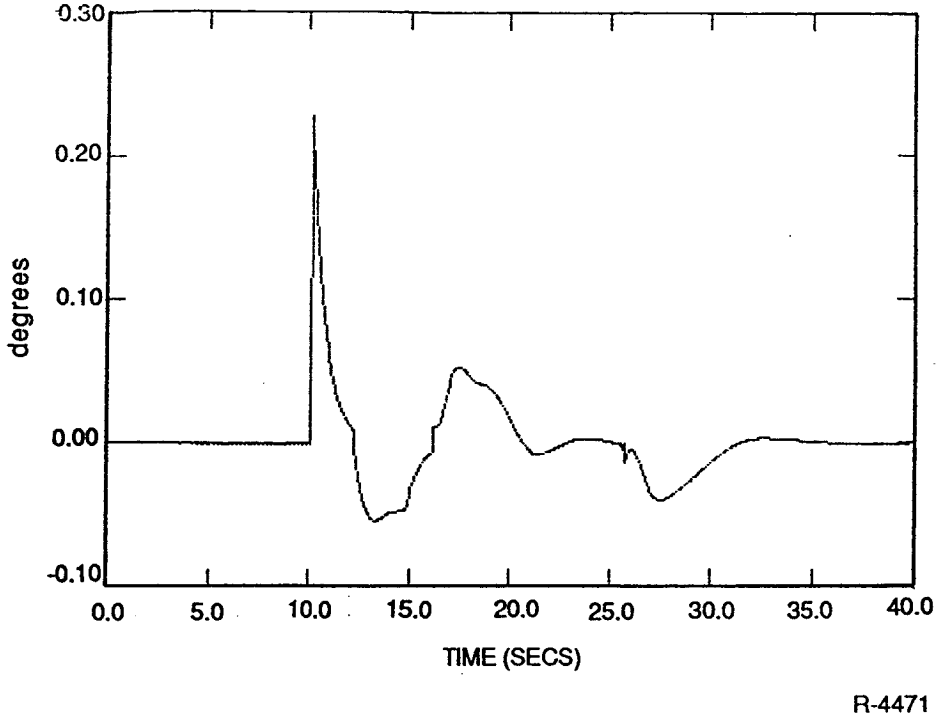


Figure 5-60. LS RES

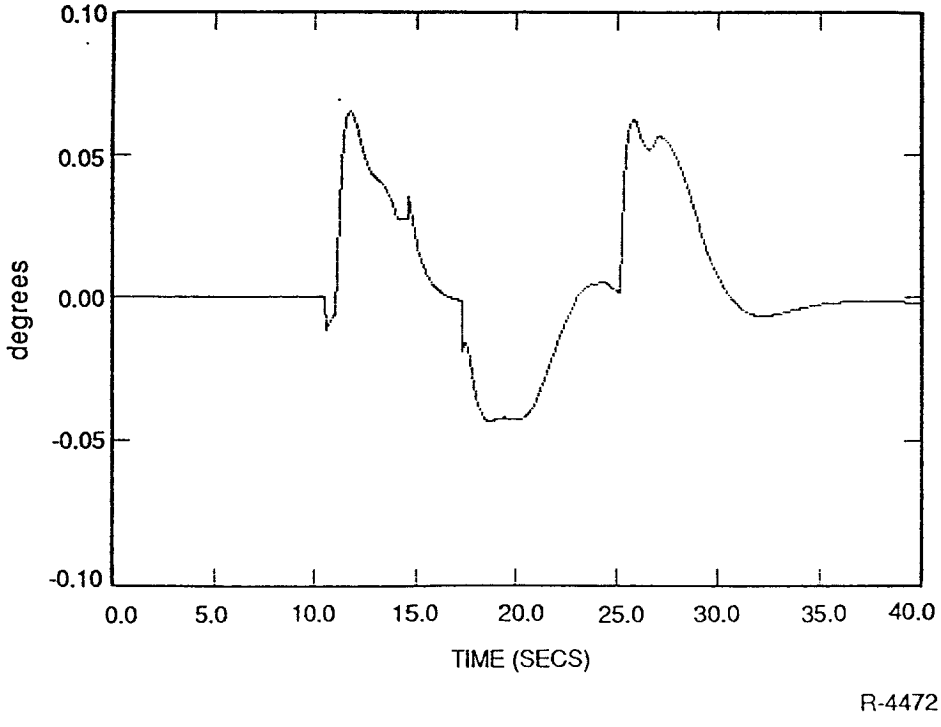


Figure 5-61. RS RES

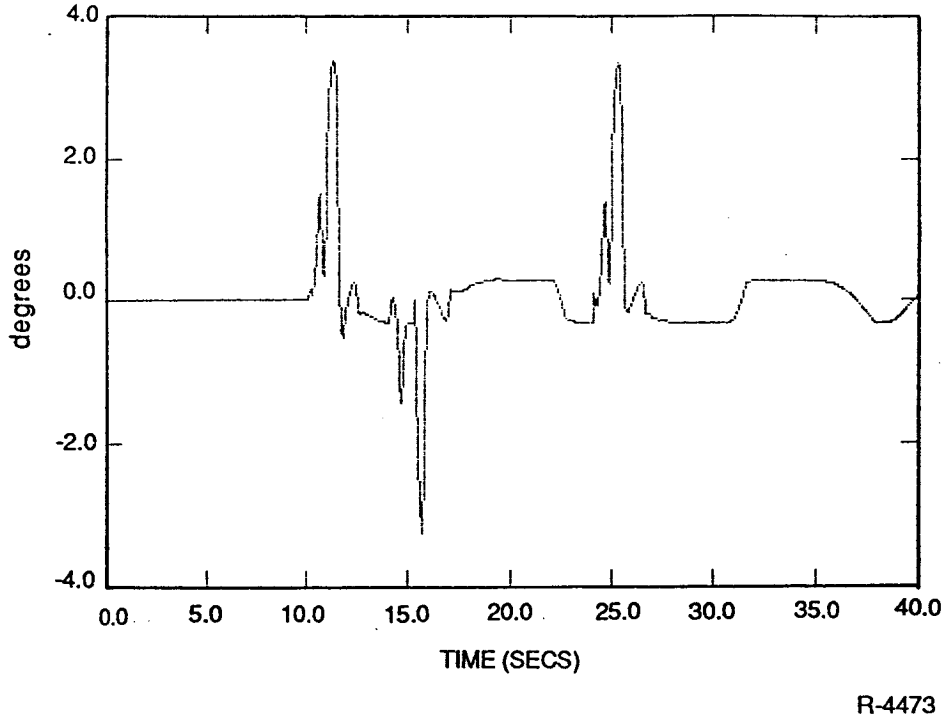


Figure 5-62. R RES

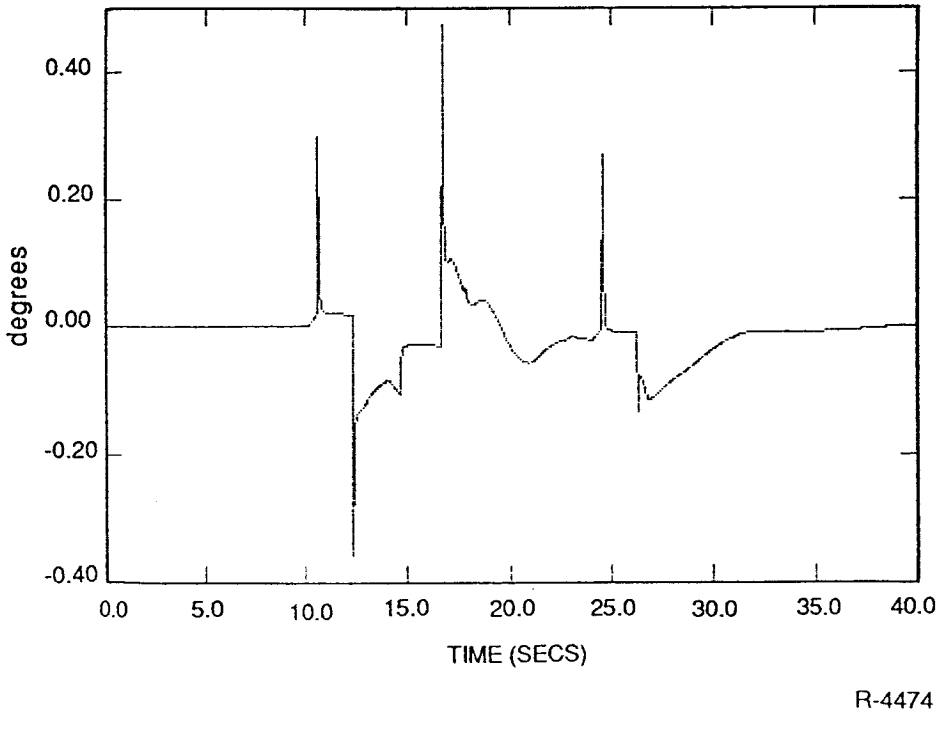


Figure 5-63. LE RES

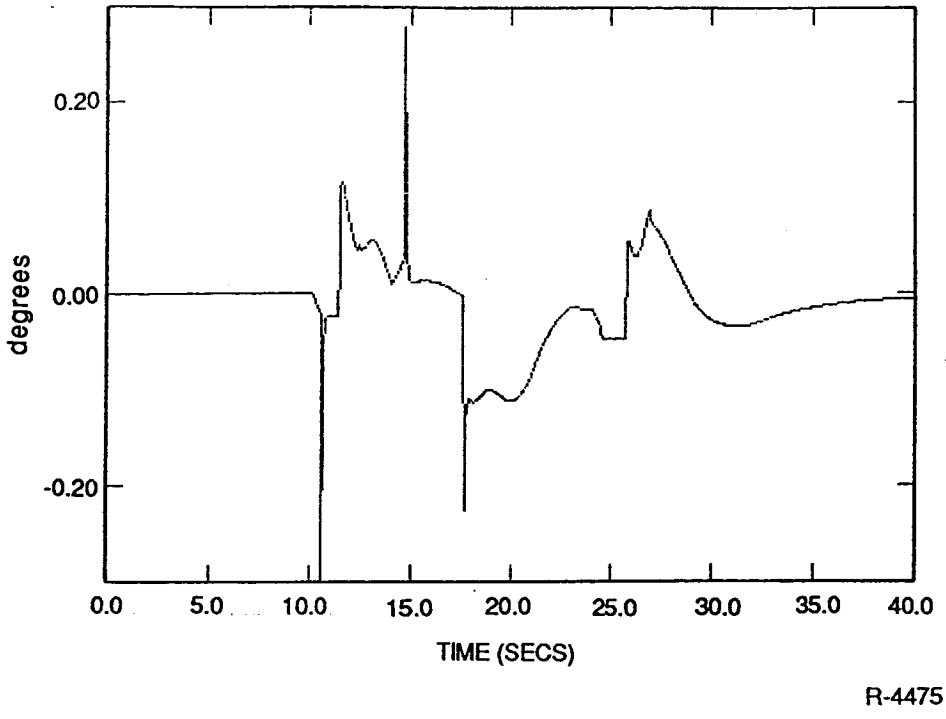


Figure 5-64. RE RES

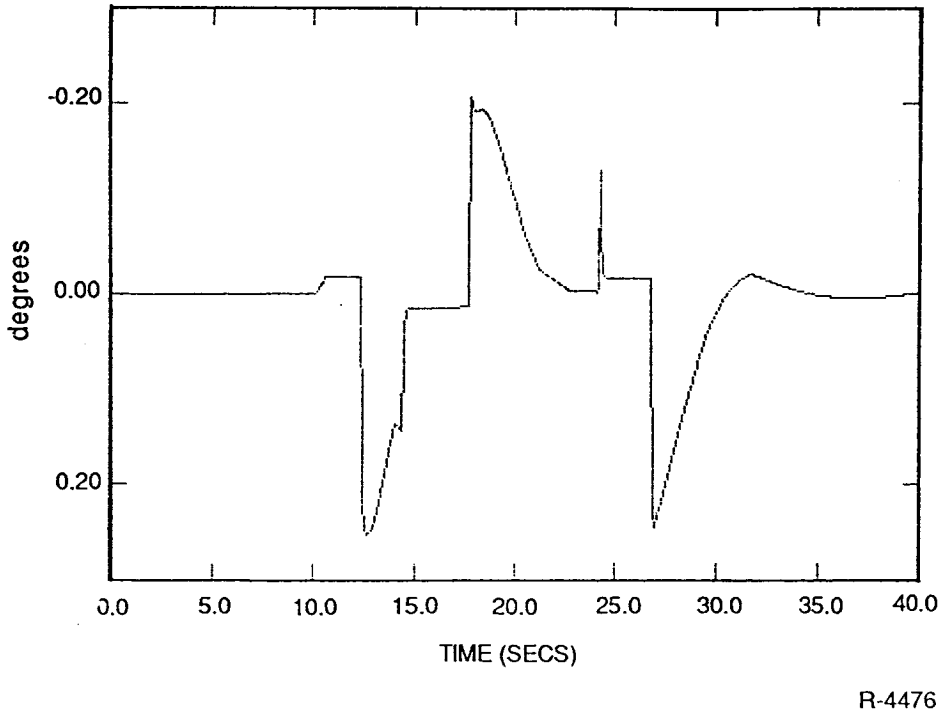


Figure 5-65. AL RES

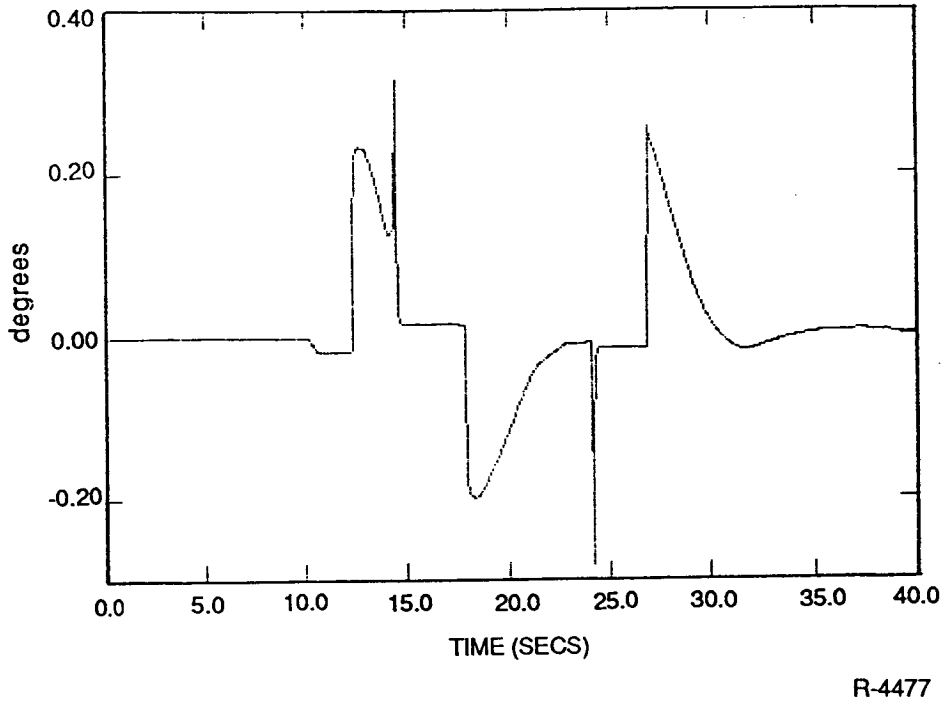


Figure 5-66. RA RES

These results motivated the development of the two decision processes discussed at the beginning of subsection 5.2. The primary focus for this project was on test and validation of a fixed threshold algorithm for every control element. Time permitted only preliminary investigation of the threshold scheduling algorithm for the aileron controls.

5.2.2 Fixed Threshold Actuator Decision Process

A structure for a fixed threshold actuator FDI algorithm is shown in Fig. 5-67. This structure implements the trigger/verify procedure described in Section 3 for dealing with unknown onset times. As in the aircraft-path FDI decision process, a "prewhitening" filter which assumes that the primary error sources are white noise and very low frequency error is employed. This filter is a high pass filter whose cutoff frequency is determined by the length of

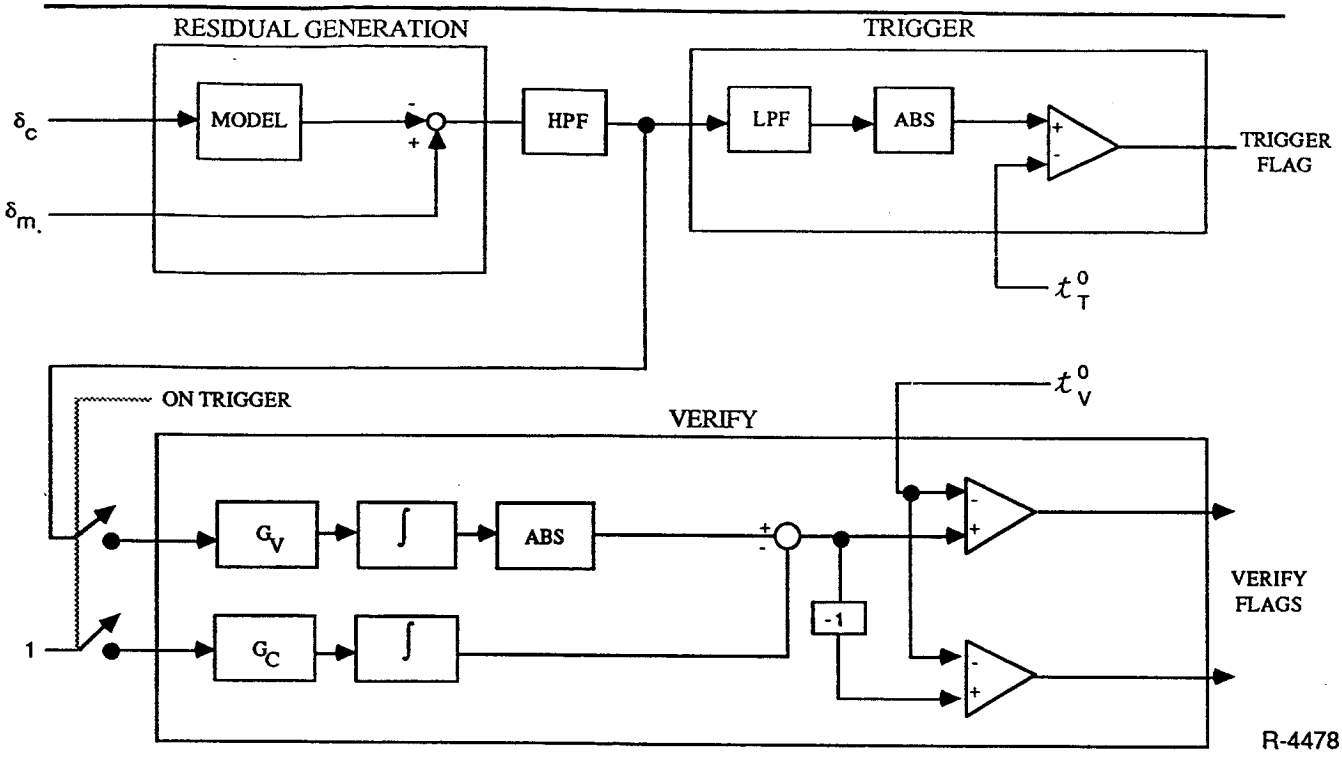


Figure 5-67. Fixed Threshold Actuator FDI Structure

time failure signatures are expected to remain constant, the length of time expected to be needed for failure detection (trigger and verify), and the bandwidth of the very low frequency errors. We assume that about one second is needed for failure detection, that the failure signatures of importance are coherent over this interval, and that the very low frequency errors have bandwidths of less than 0.1 r/s. The high pass filter bandwidth is selected at 0.5 r/s (i.e., between .1 r/s and 1 r/s). The HPF is implemented with a first order digital filter defined by

$$1 - \frac{(1-a)z^{-1}}{1-az^{-1}} \qquad (5\text{-}73)$$

where $a = e^{-0.5\Delta}$, and where z^{-1} is a unit delay.

The trigger process is based on the log-likelihood ratio test for distinguishing a bias of unknown sign in white noise from white noise. The trigger statistic is,

$$S_T(k) = \frac{1}{N} \left| \sum_{j=1}^{N} \nu(k-j+1) \right| \qquad (5\text{-}74)$$

The trigger design procedure is as follows. First, N is chosen as $1/\Delta$ times the desired trigger delay (0.5 second --> N = 10). This choice will ensure adequate averaging of high frequency errors while not reducing the effect of failures. Next, the threshold is set so that the likelihood of false alarms, during worse case errors, is small. Worst case errors, E_{wc}, were determined from simulation runs using the three doublet maneuvers, and are shown for each control in Table 5-16. If the white noise on each control measurement is 0.2 degrees (twice the allotted size in Table 5-15 for safety; this does not impact the design a great deal), then the thresholds are selected from

$$t_o = E_{wc} + 3\sigma/\sqrt{N} \qquad (5\text{-}75)$$

where σ is the .2 degree noise standard deviation. The results are shown in Table 5-17.

TABLE 5-16. WORST CASE ERRORS

SURFACE	
L&R Stabilizer	0.5 deg
L&R Elevator	0.75 deg
Rudder	3.0 deg
L&R Aileron	0.50 deg
L&R Throttle	1.5 klbs

The verify process is based on the SPRT for the alternate hypotheses defining the trigger test. The verify statistic is,

$$S_V(k) = \left| \sum_{j=k_f}^{k} G_V \, \nu(j) \right| - \sum_{j=k_f}^{k} G_c \qquad (5\text{-}76)$$

TABLE 5-17. TRIGGER THRESHOLDS

SURFACE	
L&R Stabilizer	0.6 deg
L&R Elevator	0.85 deg
Rudder	3.1 deg
L&R Aileron	0.60 deg
L&R Throttle	1.6 klbs

The test is run until $k = k_f + N_V$. The time limit N_V is taken to be 10 samples (0.5 seconds). This is deemed long enough to accomplish averaging of high frequency errors and short enough so that failure signatures are not substantially reduced. The parameter G_V is chosen as one and G_C is one half the minimally detectable failure signature, m/2. The minimally detectable signature is taken to be twice the worst case error. Finally, the threshold is determined so that false triggers are frequently rejected by $N_V/2$ and minimally detectable failures are frequently detected by $N_V/2$. This is accomplished by choosing the threshold as the expected value of $S_V(N_V/2)$. Table 5-18 shows the selected thresholds.

TABLE 5-18. VERIFY THRESHOLDS

SURFACE	
L&R Stabilizer	2.5 deg
L&R Elevator	3.75 deg
Rudder	15. deg
L&R Aileron	2.5 deg
L&R Throttle	7.5 klbs

False alarm checks were made by processing the residuals during several maneuvers with no failures according to the decision process described above. The three doublet maneuvers and the climbing turn maneuver with noise added to the sensors before processing resulted in no false triggers as expected. The climbing turn maneuver with turbulence and sensor noise resulted in false triggers of the left and right ailerons but the verify processes were able to reject these triggers as false in each case. The trigger and verify tests for the left aileron are shown in Fig. 5-68.

Detection checks were then made using the climbing turn maneuver and stuck failures (i.e., stuck at the position at the time of failure) for each control element. Unfortunately, only runs with turbulence and sensor noise were made. Since turbulence tends to excite the controls, the detection performance discussed below is presumably better than that which would be obtained without turbulence in the simulation. The failure is implemented at 5.0 seconds and the maneuver occurs at ten seconds. Table 5-19 shows the times at which the stuck failures are triggered and verified. The label (first) indicates that one or more false triggers were indicated by the verify process before a trigger is verified. Throttle failures are not detected because the deviation of the measured EPR from the estimated EPR (in klbs) does not exceed the minimally detectable level. Only one trigger occurs for the left throttle, but the signature is not large enough to be verified. This is shown in Fig. 5-69. The right elevator failure is of some interest since several triggers occur before the failure is verified. This is shown in Fig. 5-70. For the first trigger the signature is so small that the verify SPRT passes its negative threshold. For the second trigger, the verify time limit is reached and the failure is identified on the third trigger.

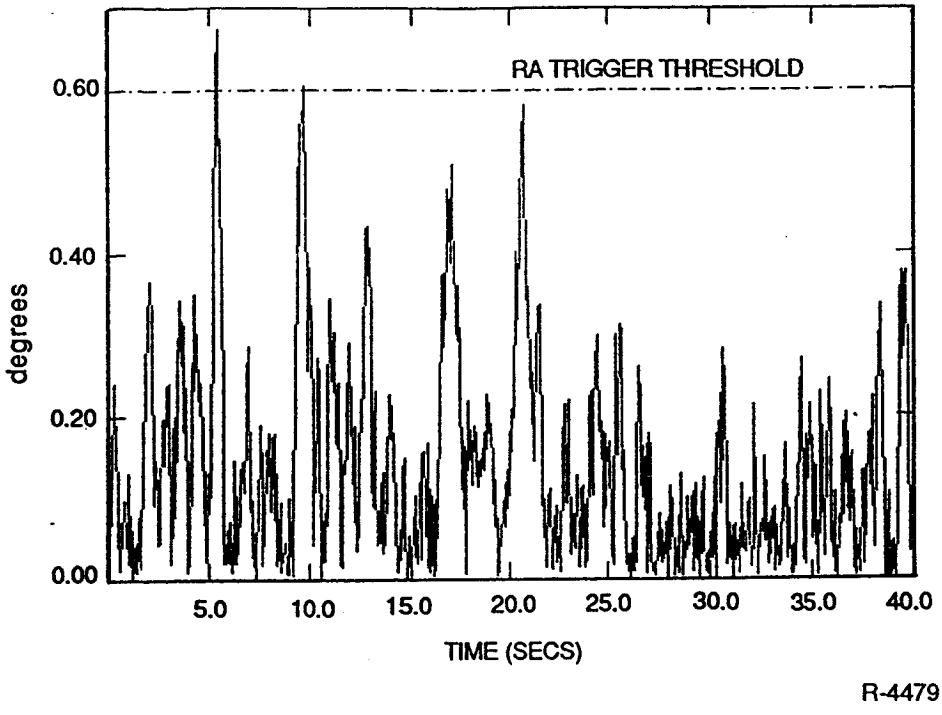


Figure 5-68a. RA TRIG

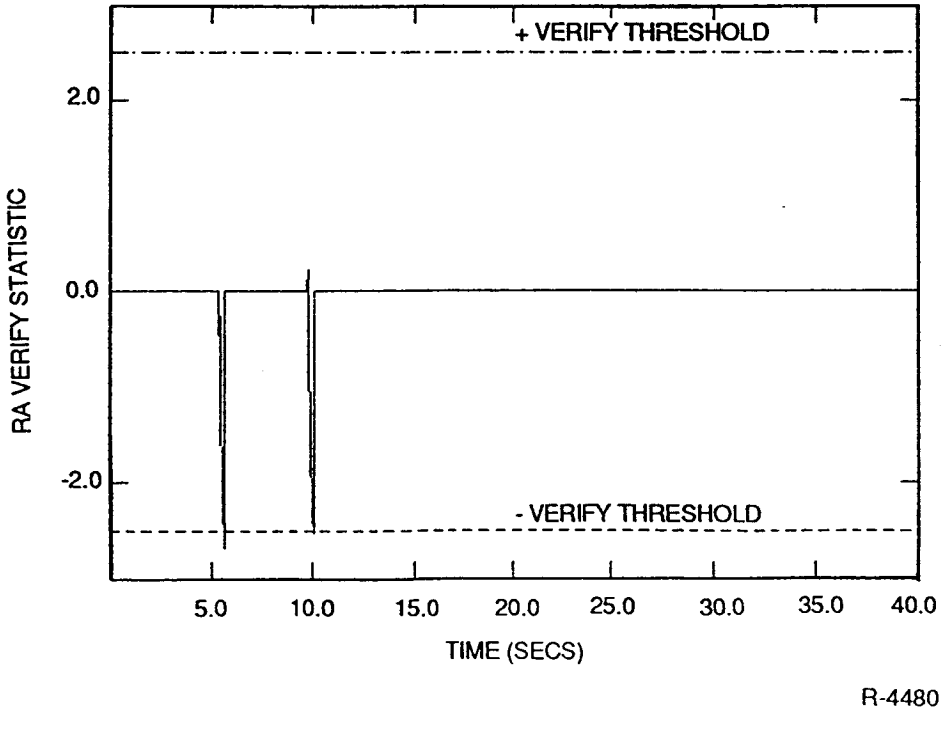


Figure 5-68b. RA VERIFY

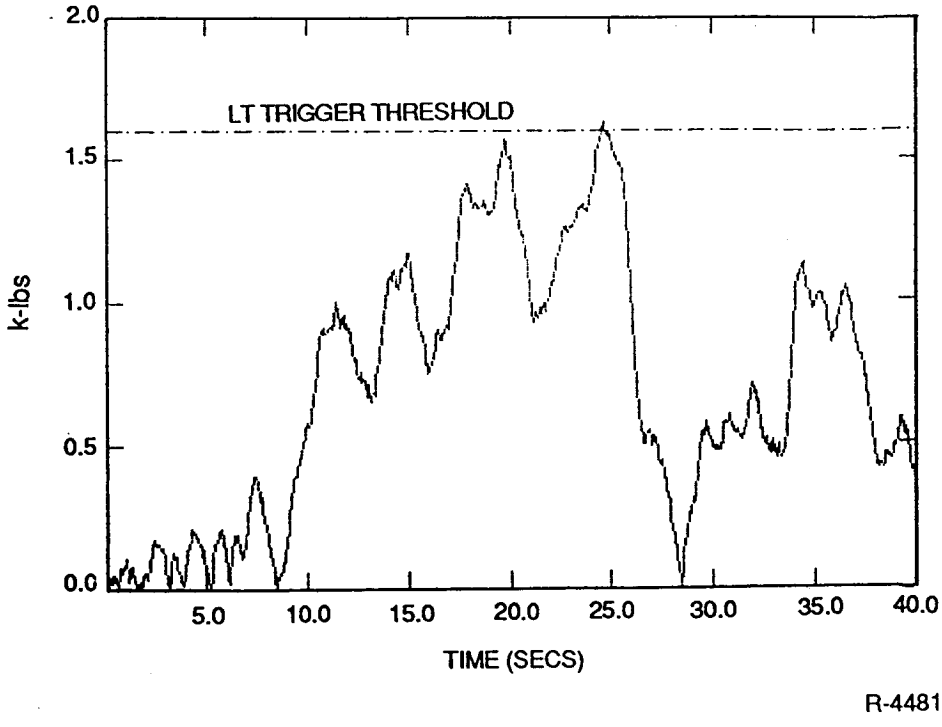


Figure 5-69a. LT TRIG

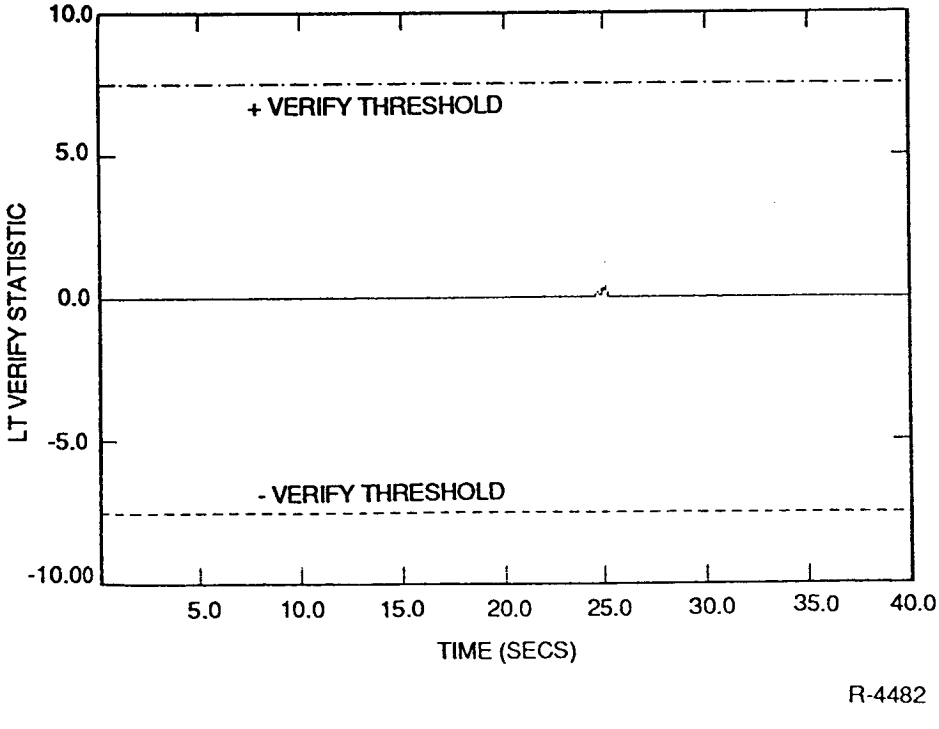


Figure 5-69b. LT VERIFY

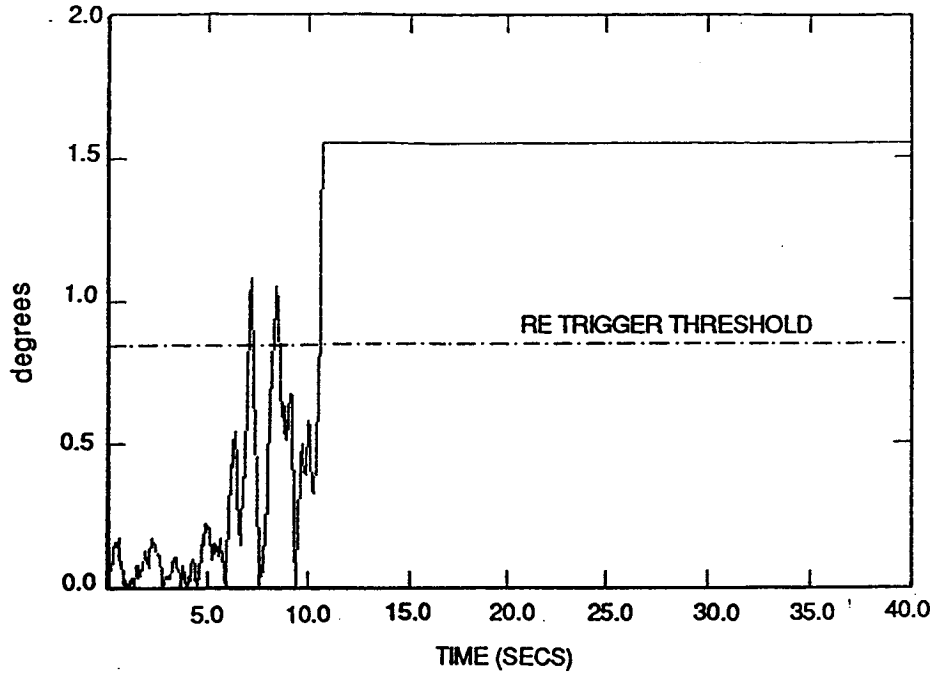


Figure 5-70a. RE TRIG

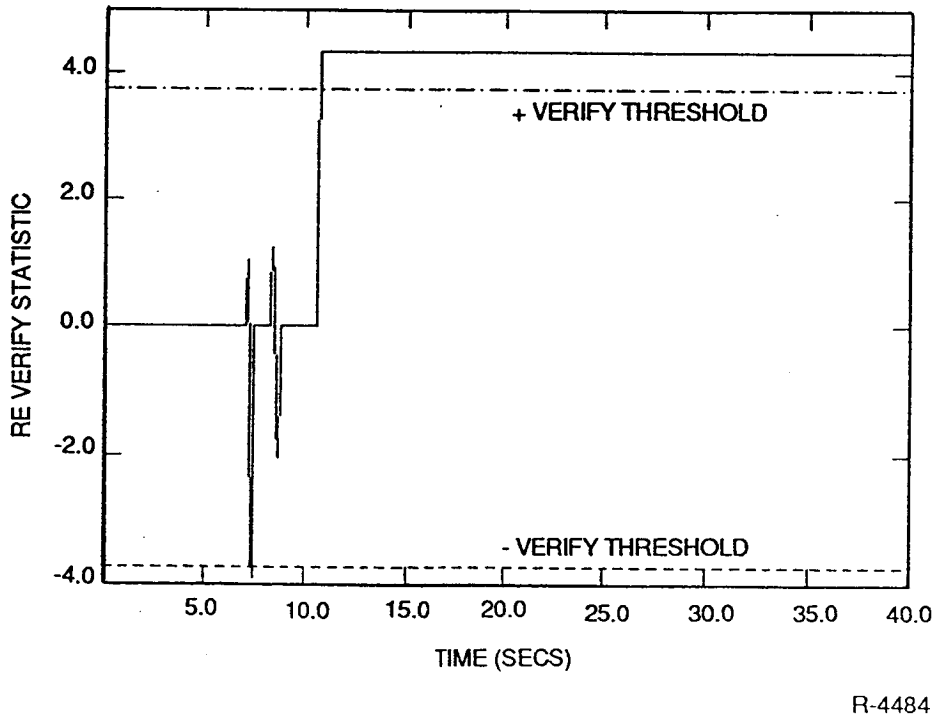


Figure 5-70b. RE VERIFY

TABLE 5-19. TRIGGER AND VERIFY TIMES FOR STUCK CONTROLS

SURFACE	TRIGGER TIME (SEC)	VERIFY TIME (SEC)
Left Throttle	24.6	---
Right Throttle	None	None
Left Stabilizer	11.25	11.5
Right Stabilizer	10.9	11.1
Rudder	6.15	6.3
Left Elevator	6.05	6.3
Right Elevator	7.0 (first)	10.7
Left Aileron	5.45 (first)	6.4
Right Aileron	5.95	6.1

The fact that throttle failures were not detected must be weighed against the severity of the failure. Inspection of the aircraft response for both failed and unfailed cases shows that almost no performance is lost during this maneuver when a single throttle is stuck. The difference appears in the airspeed in which the lowest airspeed achieved during the maneuver is five knots lower with the throttle failure. Otherwise the force and moment imbalances resulting from the throttle failure seem to be adequately compensated by the nominal FCS of [58].

DISCUSSION

The simulation results indicate that the actuator path FDI process described by Fig. 5-67 can be designed to virtually eliminate false alarms during realistic and severe maneuvers. However, the design relied on observation of the size of normal errors. Estimation or bounding of such errors is not an easy task for real aircraft. Thus, extensive flight records may need

to be examined to determine the size of errors. If these errors are expected to be constant over the flight envelope and over time, and if they are small enough (as in this study) so that important failures can be detected, then the system described above can be made to perform adequately. Otherwise, scheduling and/or adaptation mechanisms may be required.

5.2.3 Scheduled Threshold Actuator FDI Decision Process

One of the limitations of the fixed threshold algorithm is that the thresholds must take into account worst case error. This can sometimes limit the sensitivity to failures during times when this error is not likely to appear. The decision process developed here utilizes the single-input, single-output threshold scheduling ideas developed in subsection 3.3 to adjust thresholds when errors (due to dynamic model errors) are expected to be large. Time permitted only application to the aileron actuator system. These were chosen since the residuals suggested that dynamic errors were of greatest importance for these actuators.

The structure for the alternate actuator decision process is shown in Fig. 5-71. The trigger is based on the development in subsection 3.3. It is assumed that an error filter exists whose squared magnitude bounds the (relative) frequency dependent errors in the residuals. That is, the error filter in Fig. 5-70, $\ell m(\omega)$, must bound $E(\omega)/H(w)$ where $E(\omega)$ is defined in Eq. 3-125, and $H/\omega)$ is the true actuator transfer function. The trigger equations are,

$$S_T(k) = \frac{1}{N_T} \sum_{j=1}^{N_T} \nu(k-j+1)^2 \qquad (5\text{-}77)$$

$$t_T(k) = t_o + \frac{1}{N_T} \sum_{j=1}^{N_T} e(k-j+1)^2 \qquad (5\text{-}78)$$

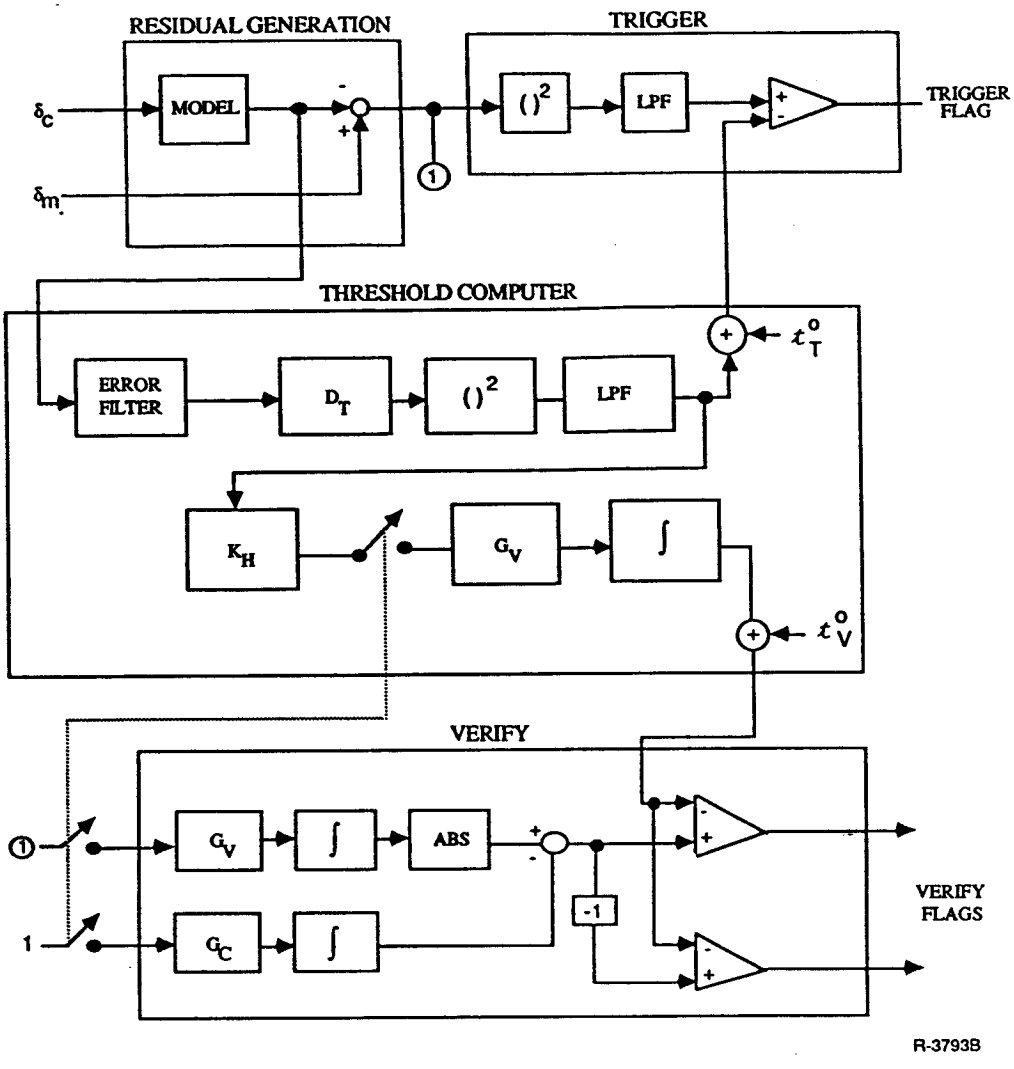


Figure 5-71. Scheduled Threshold Actuator FDI Structure
LPF = (Low Pass Filter); D_T = (Constant Multiplier)

The trigger window length, N_T, must be chosen as long as possible to ensure that the impact of the approximation of Eq. 3-128 by a finite sum is not too severe, but should not be longer than the desired FDI delay if possible. Since the time constant of the aileron actuator model is small, N_T is chosen as 10 (0.5 sec.). This will permit substantial averaging of the high frequency content in the residuals. The error filter is a first order HPF with break frequency corresponding to the aileron actuator model bandwidth. This

is based on the assumption that the first order actuator model is inaccurate above its break frequency. The nominal threshold, t_o, is now selected only to account for sensor noise since we assume that all model error will be accounted for by the threshold schedules. Thus, if $\nu(k)$ is white noise with intensity of .1 degrees, then the variance of S_T is about .02 degrees. We choose t_o as three times this or .06 degrees. Notice that this is substantially smaller than the fixed threshold algorithm. Thus, very small actuator failures can be detected during periods of quiescent operation (low frequency actuator commands) in which the threshold offsets will be small.

The verify process is based on the assumption that the nonconstant part of Eq. 5-78 is an estimate of the size of the residual error (due to dynamic mismatches) at time k. This estimate is used to modify the verify thresholds to maintain the likelihoods of error at each stage of the process as described in subsection 3.3.

The constant K_H is used to modify the high frequency gain of the error filter. The verify statistic is defined by the constants G_V and G_C. As in the trigger process G_V is chosen to be one and G_C is one half of the minimally detectable failure signature. Without errors, this signature is quite small. We chose, however, to use 0.5 degrees as a minimally desirable failure signature since smaller signatures are deemed to be unimportant and since other errors besides those accounted for by threshold scheduling may be present. The nominal threshold is chosen as in subsection 5.2.1 and takes the value of 1.25.

Several experimental design iterations were made to determine a suitable value for K_H. Choosing K_H so that the error filter has a high frequency gain of 1 (i.e., 100% error relative to the surface estimate) resulted in no

triggers during doublet and climbing turn maneuvers with no turbulence and post simulation added sensor noise. False triggers (no false alarms) occurred during the climbing turn with turbulence. Figure 5-72 shows the aileron trigger for the climbing turn maneuver. The schedules seem to be conservative (too large) but occur at the proper times. Figure 5-73 shows the same test when turbulence is added. The large high frequency excitation of the control surface results in errors which appear to be larger than those accounted for with K_H chosen to make the high frequency gain of the error filter equal to 1. It was determined that the high frequency gain of the error filter needed to be ten ($k_H = 7$) to prevent triggers during the climbing turn with turbulence. Figure 5-74 shows the resulting trigger test. The threshold offset for this test is frequently larger than required. Although it is believed that large relative error is indeed possible for some frequencies, it is likely that it is not constant over a large frequency range as in the error filter. This may explain the over conservativeness of the test shown in Fig. 5-74. An alternate method which utilizes absolute error can easily be derived (amounting to using the actuation command instead of the estimate in the error filter) and may produce better results.

Finally, detection checks were made using simulations of stuck actuators during the climbing turn maneuver with turbulence and sensor noise. Although failures are easily detected, no significant improvement in detection time (over the fixed threshold case) was observed. Figure 5-75 shows the trigger and verify tests for this case.

DISCUSSION

It is beieved that more work on the threshold scheduling method is needed before meaningful conclusions can be drawn.

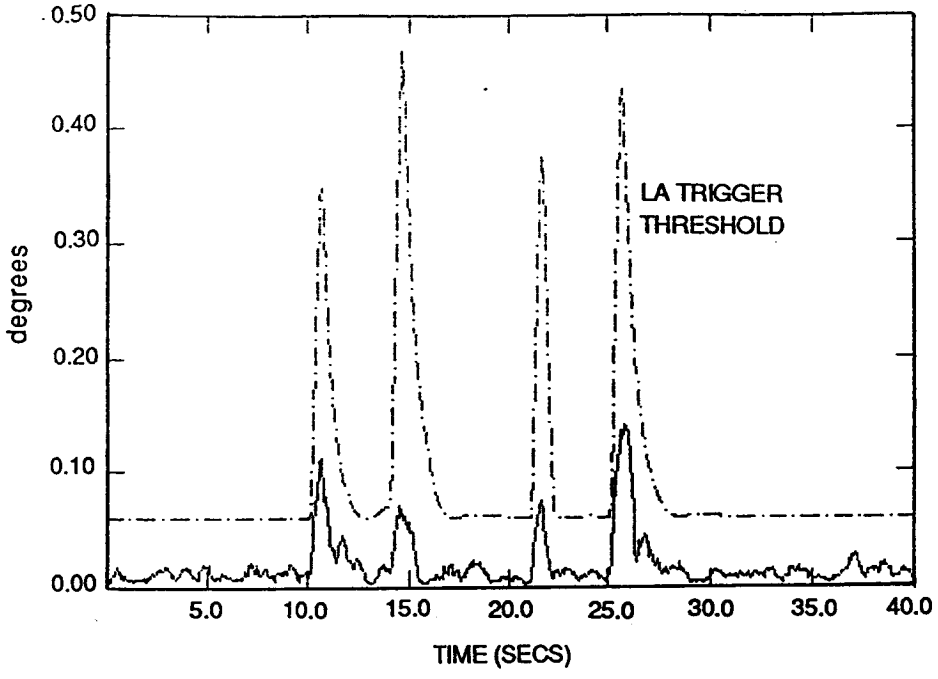


Figure 5-72. LA TRIG (—— = Statistic) (-•-• = Threshold)

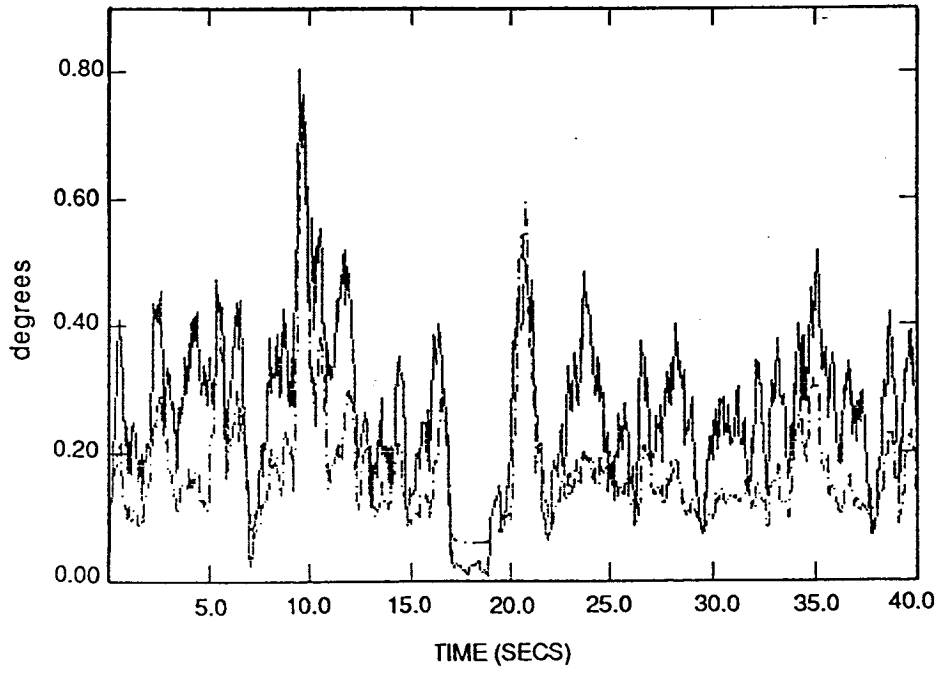


Figure 5-73. LA TRIG (—— = Statistic) (-•-• = Threshold)

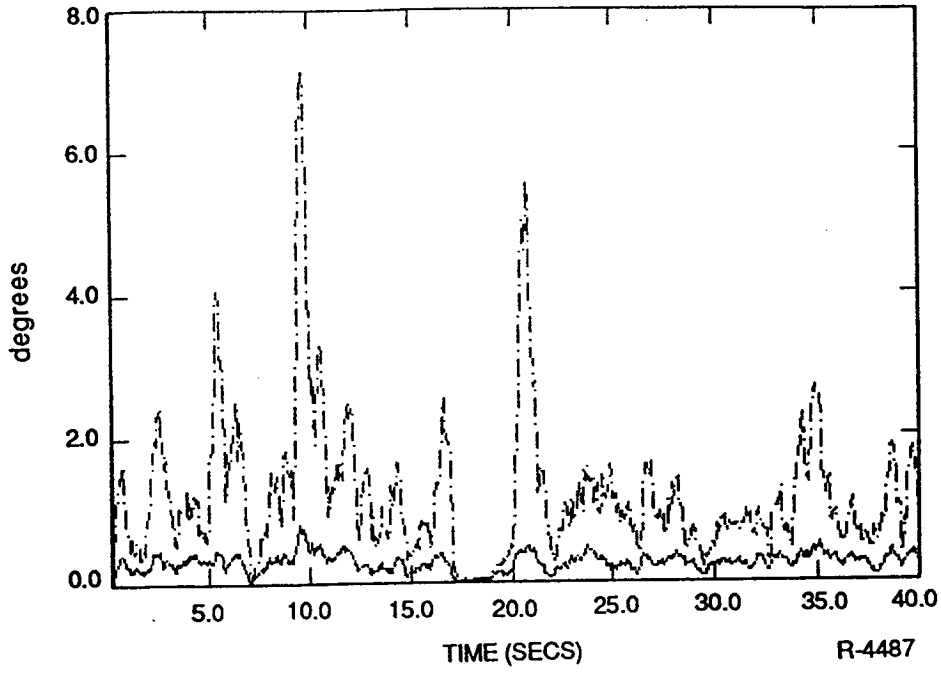


Figure 5-74. LA TRIG (KH = 7) (——— = Statistic) (-•-• = Threshold)

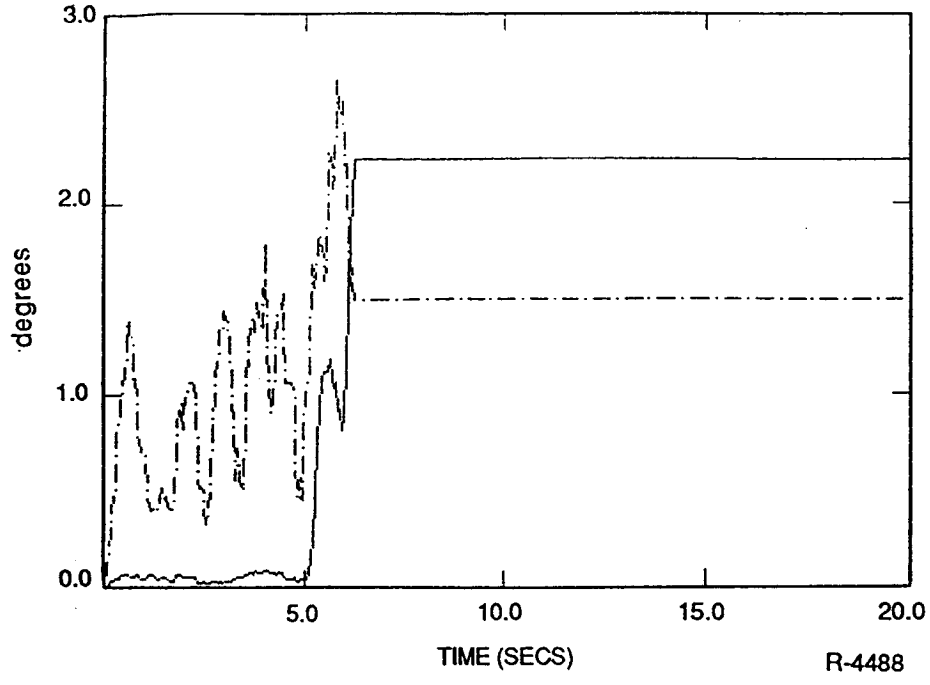


Figure 5-75a. RA TRIG (—— = Statistic) (-•-• = Threshold)

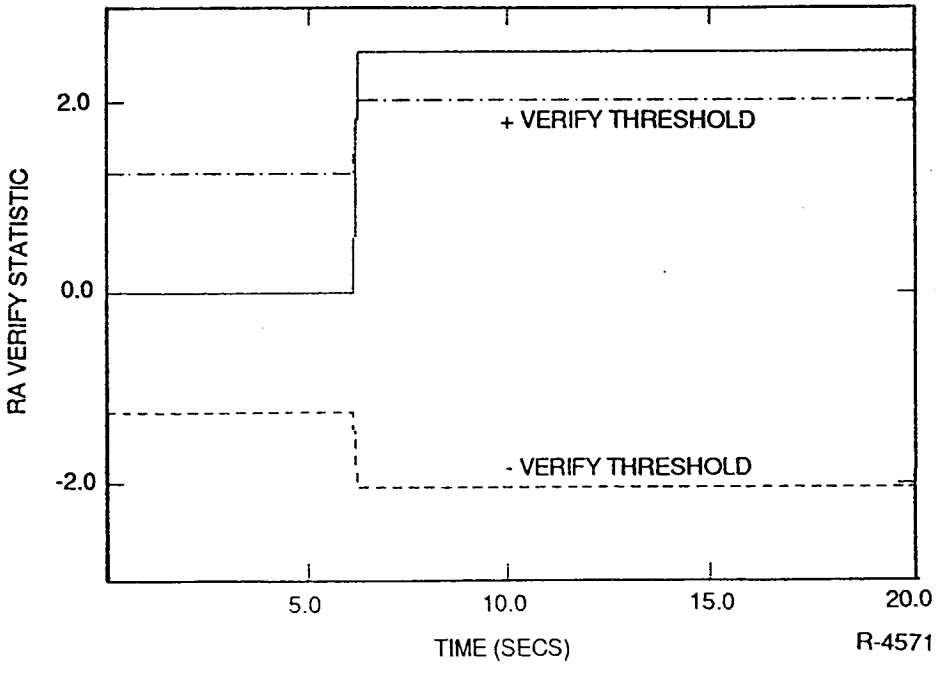


Figure 5-75b. RA VERIFY (—— = Statistic) (-•-• = Threshold)

SECTION 6

SUMMARY OF RESULTS AND CONCLUSIONS

This effort has developed and explored the use of a decentralized approach to failure detection and isolation for use in restructurable control systems. This work has produced,

1. A method for evaluating fundamental limits to FDI performance,
2. Application of this method using flight recorded data,
3. A working control element FDI system with maximal sensitivity to critical control element failures,
4. Extensive testing of this system on realistic simulations, and
5. A detailed design methodology for this system which involves parameter optimization (with respect to model uncertainties) and sensitivity analyses.

For this project, we have concentrated on detection and isolation of generic control element failures since these failures frequently lead to emergency conditions and since knowledge of remaining control authority is essential for control system redesign. The failures we considered are generic in the sense that no temporal failure signature information was assumed. Thus, various forms of "functional failures" are treated in a unified fashion. Such a treatment results in a robust FDI system (i.e., one that covers all failure modes) but sacrifices some performance when detailed failure signature information is known, useful, and employed properly. We assumed throughout this project that all sensors are validated (i.e., contain only in-spec errors) and that only the first failure of a single control element needs to be detected

and isolated. The FDI system which has been developed will handle a class of multiple failures.

The FDI system which was developed using the design methodologies outlined in Section 3 worked quite well during simulation tests on data from NASA's modified B-737 simulation. This is true despite large errors between the models used in the FDI system and the models used in the simulation. Furthermore, this system worked as predicted when errors between the simulation model and the FDI models were within the envelope used to choose FDI parameters (thresholds, etc.). The design methodology discovered the inherent indistinguishability of same-side elevator and stabilator panels (on the basis of force and moment imbalances alone) early in the project resulting in a system which does not attempt to do what is impossible.

Improvements in the design methodology, system detail, and implementation are still possible. Some over-conservatism was observed in the design procedures for sequential tests. Alternative filter structures may speed up the FDI process. The use of "probe" signals during the isolation phase would allow enhanced isolation capability, (including the isolation of stabilizers from elevators). And, finally, further exploration of the importance and design of threshold schedules is needed.

The remainder of this section provides a more complete summary of this report. It also provides a discussion of some of the unexplored alternatives in the design, implementation, and testing of the FDI system, and suggestions for further work.

DECENTRALIZED FDI

In Section 3 of this report, ALPHATECH's decentralized approach to FDI was described. This approach requires an assessment of all sources of

redundancy (including analytical redundancy) and the utilization of this redundancy to produce decoupled or decentralized residual signals that are then processed to produce FDI decisions. This method represents a suboptimal approach to the data fusion problem (i.e., combining all sources of information for decisionmaking purposes) under ideal circumstances (no modeling error), but prevents the mixing of well known relationships with poorly known relationships when modeling error exists. As a result, the decentralized approach can be superior when significant modeling errors exist.

The first application of this idea to the control element FDI problem resulted in a decomposition into actuator-path and aircraft-path subproblems. Measurements of actuator inputs and outputs allowed this decomposition. The advantage of this decomposition is that uncertainties about aerodynamic models do not affect the ability to detect actuator path failures and uncertainties about the actuation mechanism do not affect the detectability of aircraft path failures. The disadvantage of this decomposition is that some configurations of actuator output sensors may be vulnerable to failure thereby complicating the (assumed) sensor validation process. Fortunately, however, the distinction between the aircraft-path and the actuator-path is somewhat arbitrary so that the aircraft path subsystem can easily be designed to detect actuator failures when actuation outputs are not available.

The second application of decentralization is in the method of forming aircraft-path residuals. Since only force and moment balance relationships are affected by aircraft-path failures, residuals based on these relationships are desired. Translational accelerometer measurements allows this to be accomplished for the force balances, however, rotational dynamic relationships must also be used since rotational acceleration is not measured. This method

of forming aircraft path residuals has two advantages. The first is that errors in other relationships which would affect centralized (Kalman filter-like) residual generators do not affect these decentralized residuals. A prime example is the effect of acceleration of the airmass on the translational kinematic relations. The second advantage is that control element failures show up in fixed directions in residual space. This permits the distinguishability of aircraft-path control element failures independent of the detailed failure signatures. Only the relative magnitude of imbalances in different residuals is necessary for failure isolation.

FDI SYSTEM

The structure of a generic FDI system which efficiently solves the unknown onset time problem was developed in subsection 3.1. This structure involved a monitoring or trigger process which is used to reject the hypothesis of normal operation and to trigger a verification and isolation process to reject false triggers and to identify the source of a failure. This structure is used to achieve performance advantages which approach the performance of the known onset-time case. These advantages include greater failure sensitivity, lower false alarm rates, and shorter detection delays.

Aircraft-Path Subsystem - The aircraft-path trigger was designed to make the probability of missing a critical failure small. Thus, each failure mode had an explicit trigger function which is optimized for triggering under the corresponding failure mode. Each trigger satisfies the condition that IF a particular minimal failure occurs, THEN the corresponding trigger test will "pass." Since the converse is not true and since false triggers are possible, we need to perform verify and isolate tests.

The verify and isolate tests are sequential tests and are designed so that failures which are larger than some minimal value will be detected and isolated in shorter time periods. If they reach a maximal time limit, no decision is made, although it is also possible to make a fixed sample size decision at this time.

The isolation process recognizes the fact that only the rejection of failure modes is possible when signature information is not used. This fact results in principle, in a matrix of isolation tests, each designed to reject a failure mode with maximal sensitivity to another failure mode. Although this structure appears complex, it guarantees optimal performance for every failure mode and allows detailed analysis and optimization of each part of the system. In practice, the off-diagonal tests in this isolation matrix were combined for efficiency. Also, in principle, only those failure modes which are in the "trigger-implied ambiguity group" need to be isolated, although in practice all failures were consdiered as possible following any trigger. To declare a failure, all isolation tests must "vote" in favor of that failure.

Actuator-Path Subsystem - The character of the actuator residuals (all actuator failure directions are mutually orthogonal) resulted in one actuator-path subsystem for each actuator failure. Thus, no isolation process was needed. These subsystems, like the aircraft-path subsystem, also used a trigger/verify structure to "solve" the unknown onset time problem. Two decision processes were created and tested; a fixed threshold and a varying threshold algorithm.

The fixed threshold algorithm was designed to accommodate the observed low frequency behavior in each residual, sensor noise, and other high frequency

errors. The result of a trigger crossing its threshold is the initiation of a sequential verify test. If the verify test passes, the corresponding control element is declared as failed. If a verify fails, a "false trigger" is declared. Because fixed thresholds were used to accommodate low frequency errors, the sensitivity to actuator path failures was higher than originally expected (though by no means unacceptable).

The varying threshold algorithm was based on the concept derived in Section 3 for single-input, single-output systems with transfer function errors. It assumed that all transfer function errors were high frequency relative errors. Observations clearly indicated that this was not the case, and consequently, this decision process did not perform as well as expected. Further work in this area is needed before substantive conclusions can be drawn.

DESIGN METHODOLOGIES

One key to the successful deployment of any FDI concept (or any aircraft system concept for that matter) is the development of an analytical design methodology which permits engineers to assess the impact of various contingencies on performance and modify the design accordingly. In Section 3 of this report, we have outlined the structure of such a methodology and have given examples of how the many analysis and synthesis tasks could be performed. The reason such a methodology is important is that exhaustive simulation and flight testing for purposes of design is expensive and may never encompass all contingencies of interest. An analytical design method provides quicker and cheaper answers to the same questions and should allow all contingencies to be considered.

For the FDI problem, the contingencies of interest are the size of potential error sources and the magnitude and character of the maneuvers which excite them. In Section 4 we demonstrated how an error budget might be formed and utilized in the design process. This error budget serves to define a truth model for which parameters in the FDI system (which is designed on the basis of a design model) are optimized.

SIMULATION RESULTS

Substantial testing of both subsystems was performed using data from NASA's B-737 simulation. These tests included drastic model mismatches between the FDI models and the actual simulation model, sensor noise and scale factors and turbulence. Various combinations of sensor noise and turbulence were used to evaluate the impact of each on the FDI system. A total of 41 60-second simulation runs were made by NASA and data recorded for use in this project. All runs were made using the control law of [58].

Aircraft-Path Subsystem - Three categories of simulation runs were used to test this subsystem. These included:

1. False alarm checks with three "doublet" maneuvers and a climbing turn. The climbing turn checks were made with no turbulence or sensor noise, with sensor noise alone, and with both sensor noise and turbulence. The doublet maneuver checks were made with no sensor noise and no turbulence and with sensor noise and no turbulence.

2. Detection checks were made using 100 percent effectiveness failures of all control elements except the engine (simulation capability for this case was not available) during a climbing turn maneuver. These checks were made with sensor noise alone--no turbulence.

3. Detection checks were made for varying levels of partial effectiveness failures for the left aileron during a climbing turn with sensor noise and no turbulence.

For the nine false alarm check runs which were made, no false alarm was ever generated. No false triggers occurred during any of the three climbing turn cases. During the two pitch doublet maneuvers triggers of the left "horizontal tail" control occur, but are verified as false. During the two roll doublet maneuvers, triggers of every control element occur, but are also verified as false.

Of the seven control elements which were failed 100 percent in the second category of simulation runs, five were correctly isolated within two seconds of either the failure time (for surfaces with loads in straight and level flight) or the maneuver time. Left and right elevator failures continually caused triggers to occur, but no unanimous decision could ever be reached during the climbing turn maneuver. Fortunately (and as expected), the degradation in maneuver performance due to these failures is slight and they clearly do not represent an emergency situation for the aircraft with the control law being used. In no cases was an incorrect isolation ever made. Of course, no attempt was made to isolate same side elevator and stabilizer controls from each other since they were deemed indistinguishable (using force and moment imbalances alone) during the design process. Thus correct isolation for the stabilizers implies isolation to a fictitious horizontal tail surface. The elevator failures which were not unanimously isolated turned out to have signatures which were considered only marginally detectable by the design methodology and thus this is an expected result.

Finally, of the five left aileron failures of varying partial effectiveness, three were correctly isolated. It was concluded that failures of greater magnitude than sixty percent loss could be correctly isolated during this maneuver and of less magnitude than forty percent could not be unanimously

isolated. The forty percent failure case caused triggers to occur and the twenty percent case did not. Isolation times varied from less than one second for 100 percent failures to five seconds (two triggers) for the sixty percent failure. These results are consistent with the notion that for a given maneuver, the signature magnitude decreases with failure severity, thus reducing the ability to detect and isolate and increasing the length of time needed for the sequential decision process to conclude.

No detection checks were made with turbulence, however, it would be anticipated that due to increased excitation of controls, performance would be equal or better to what was observed. Some experimentation with "forced isolate decisions" were made leading to the conclusion that, for the case tried, such a procedure may provide significantly enhanced detection performance with little degradation in false alarm performance. More discussion of the results is provided at the end of subsection 5.1.

Actuator-Path Subsystem - For the fixed threshold decision algorithm false alarm and detection of stuck-at-failure checks were made. The false alarm checks were made using a climbing turn with sensor noise and no turbulence and with both noise and turbulence and using three doublet maneuvers with sensor noise and no turbulence. Detection checks were made using the climbing turn with sensor noise and turbulence. Time did not permit examining the effects of sensor noise and turbulence independently. Thus detection performance as sited here is presumably better than what would be ovserved with no turbulence.

None of the four maneuvers with noise and no turbulence produced a false trigger. When turbulence was added during the climbing turn maneuver false

triggers of the left and right aileron were observed, but no false alarms were generated.

Of the nine stuck-at failures implemented, seven were correctly detected. Only the engine failures were missed due to the lack of sufficient excitation. Examination of the impact of missing these failures showed that little degradation in performance occurred (five knot transient airspeed difference). Failure detection times varied from 150 milliseconds to greater than three seconds. The longer detection times were due to the fact that triggers occurred in unloaded surfaces before the maneuver as a result of excitation due to turbulence.

The varying threshold algorithm was also tested, however, too few results are available to draw significant conclusions.

UNEXPLORED CONCEPTS, FURTHER WORK, AND OTHER NOTES

A list of other comments including concepts and implementations which were left unexplored, some suggestions for possible algorithm improvement, and other notes is provided below. This is in addition to the suggestions made at the end of subsections 5.1 and 5.2.

1. In the design methodology, a more detailed truth model might be used to alleviate some of the conservatism inherent in the current design. This might include modeling of the pilot and the control law to obtain more accurate statistical descriptions of the residuals. Perhaps a continuous time domain methodology would be more useful than the discrete models currently being used. Note, however, that these statistical models would still assume, at least, piecewise stationarity and that nonstationarity is a large issue which must be handled in some fashion.

2. In the design methodology, more detailed sensitivity analyses would be useful in demonstrating the tradeoffs associated with the design.

3. Computational savings and even performance improvement may be achieved by implementing the verify and isolate tests on a "trigger implied ambiguity basis." The ambiguity, of course, depends not only

on the physics of the problem, but also on the trigger design. For example, if only the yaw residual is used to trigger rudder failures, then a rudder trigger alone eliminates those controls which can not excite yaw imbalances.

4. In the design process the notion of an FDI bandwidth arose. Failures whose signatures lie outside this bandwidth will not be detectable. If it were desirable to detect such failures (especially very low frequency signatures), one might consider the design of a multiple bandwidth FDI system. Such a system would have higher detectabilities for failures which are outside the current bandwidth without sacrificing the performance obtained with the current system. Of course, the limiting case of a multiple bandwidth system would be a concept involving Fourier transforms and statistics and thresholds for combinations of frequency elements.

5. In the evaluation of fundamental limits to FDI performance, it would be useful to process more segments of flight data and to average the results. Use of other spectral estimation methods might also be useful, particularly the periodogram approach (i.e., taking averages of DFT's). These may be explored in subsequent work.

6. One dissapointment of the current work was the need to rely on observed data in the design process. Studies relating to the creation of more accurate error budgets need to be accomplished in order for the design methodology to be maximally useful.

7. A reduction in computational requirements for the aircraft path algorithm might be achieved by a system which performed only a single rejection test for each control element. Such a system would, of course, sacrifice sensitivity to failures. However, this approach might be tried first and evaluated to see where specific sensitivities need to be enhanced. The system concepts developed in this work could then be applied only where increased sensitivity was needed.

8. Relaxation of sensor validity assumptions is an important future effort. To detect sensor and control element failures, additional redundancy relations are needed. The work of Deckert, et al., [16] provides an excellent starting point since it was used as a basis for the approach developed in this work. Reliability issues also need to be addressed in this context to ensure that the tradeoff between hardware and analytic redundancy is made properly.

9. The actuator models used in this work required a substantial amount of detail (scale factors, rate limits, etc.). The most significant detail seems to have been rate limits since errors due to incorrect rate limits can affect both low and high frequency behavior.

10. For the varying threshold actuator decision process, improved performance may be possible if bounds on the absolute error (rather than the relative error) are used to schedule thresholds.

APPENDIX A

DERIVATION OF AERODYNAMIC COEFFICIENTS FROM LINEAR MODELS

Assume that the linear perturbation model was derived through a first order Taylor expansion of Euler's equations. That is, the linear model

$$\dot{x} = Ax + Bu \tag{A-1}$$

represents the dynamic perturbations from nominal values x_0 and u_0, derived by a linearization of the nonlinear rigid body equations. The state vector x is (U, V, W, P, Q, R) and the control vector u is composed of the "deflections," δ_i (including throttle). Equation A-1 was derived from a linearization of the following equations.

$$m(\dot{U} + WQ - RV) = X \tag{A-2}$$

$$m(\dot{V} + UR - PW) = Y \tag{A-3}$$

$$m(\dot{W} + VP - QU) = Z \tag{A-4}$$

$$I_x\dot{P} + (I_z - I_y)RQ = L \tag{A-5}$$

$$I_y\dot{Q} + (I_x - I_z)PR = M \tag{A-6}$$

$$I_z\dot{R} + (I_y - I_x)QP = N \tag{A-7}$$

The forces and moments (X Y Z L M N) were expressed in Eqs. 5-7 through 5-18 in terms of a measurement vector, $y^T = (V_T, \alpha, \beta, P, Q, R)$ and the nondimensional coefficient which we wish to solve for. Thus, Eqs. A-2 to A-7 can be rewritten,

$$\dot{x} = f(y, \delta ; C) \qquad \text{(A-8)}$$

where C is a vector of nondimensional coefficients. Since y is a function of x the partial derivatives $\frac{\partial \dot{x}_i}{\partial x_j}\Big|_{x_o u_o}$ and $\frac{\partial \dot{x}_i}{\partial \delta_j}\Big|_{xuU_o}$ can be obtained as a function of C (Note: $\frac{\partial \dot{x}_i}{\partial x_j} = \frac{\partial \dot{x}_i}{\partial y_j}\frac{\partial y_j}{\partial x_j}$). These partials are then assigned to the numerical values specified by the A & B matrices in Eq. A-1 and the coefficients C found.

In particular, if we assume that V_T is constant, then the solution for C is unique and is given below.

Stability Terms:

$$C_{X\alpha} = m\left[\left(\frac{\partial \dot{U}}{\partial U}\right)(-V_T \sin\alpha \cos\beta) + \left(\frac{\partial \dot{U}}{\partial W} + Q\right) V_T \cos\alpha \cos\beta\right]/\bar{Q}S \qquad \text{(A-9)}$$

$$C_{XQ} = m\left(\frac{\partial \dot{U}}{\partial Q} + W\right) 2V_T / \bar{c}\,\bar{Q}S) \qquad \text{(A-10)}$$

$$C_{Y\beta} = \left[\left(\frac{\partial \dot{V}}{\partial U} + R\right)(-V_T \cos\alpha \sin\beta) + \left(\frac{\partial \dot{V}}{\partial V}\right) V_T \cos\beta + \frac{\partial \dot{V}}{\partial W} - P\right)$$
$$(-V_T \sin\alpha \sin\beta)\Big] / \bar{Q}S \qquad \text{(A-11)}$$

$$C_{YP} = \left(\frac{\partial \dot{V}}{\partial P} - W\right)(2V_T / b\bar{Q}S) \qquad \text{(A-12)}$$

$$C_{YR} = \left(\frac{\partial \dot{V}}{\partial R} + U\right)(2V_T / b\bar{Q}S) \qquad \text{(A-13)}$$

$$C_{Z\alpha} = \left[\left(\frac{\partial \dot{W}}{\partial U} - Q\right)\left(-V_T \sin\alpha \cos\beta\right) + \left(\frac{\partial \dot{W}}{\partial W}\right) V_T \cos\alpha \cos\beta\right] / \bar{Q}S \quad \text{(A-14)}$$

$$C_{ZQ} = m\left(\frac{\partial \dot{W}}{\partial Q} - U\right)\left(2V_T / c\bar{Q}S\right) \quad \text{(A-15)}$$

$$C_{L\beta} = I_x\left[\left(\frac{\partial \dot{P}}{\partial U}\right)\left(-V_T \cos\alpha \sin\beta\right) + \left(\frac{\partial \dot{P}}{\partial V}\right) V_T \cos\beta + \frac{\partial \dot{P}}{\partial W}\right)$$
$$\left(-V_T \sin\alpha \sin\beta\right)\Big] / \bar{Q}Sb \quad \text{(A-16)}$$

$$C_{LP} = I_x\left(\frac{\partial \dot{P}}{\partial P}\right)\left(2V_T / b^2\bar{Q}S\right) \quad \text{(A-17)}$$

$$C_{LR} = \left[I_x\left(\frac{\partial \dot{P}}{\partial R}\right) + \left(I_z - I_y\right) Q\right]\left(2V_T / b^2\bar{Q}s\right) \quad \text{(A-18)}$$

$$C_{M\alpha} = I_y\left[\left(\frac{\partial \dot{Q}}{\partial U}\right)\left(-V_T \sin\alpha \cos\beta\right) + \left(\frac{\partial \dot{Q}}{\partial W}\right) V_T \cos\alpha \cos\beta\right] / \bar{Q}S\bar{c} \quad \text{(A-19)}$$

$$C_{MQ} = I_y\left(\frac{\partial \dot{Q}}{\partial Q}\right)\left(2V_T / \bar{c}^2\bar{Q}S\right) \quad \text{(A-20)}$$

$$C_{N\beta} = I_z\left[\left(\frac{\partial \dot{R}}{\partial U}\right)\left(-V_T \cos\alpha \sin\beta\right) + \left(\frac{\partial \dot{R}}{\partial V}\right) V_T \cos\beta + \left(\frac{\partial \dot{R}}{\partial W}\right)\right.$$
$$\left.\left(-V_T \sin\alpha \sin\beta\right)\right] / \bar{Q}Sb \quad \text{(A-21)}$$

$$C_{NP} = \left[I_z\left(\frac{\partial \dot{R}}{\partial P}\right) + \left(I_y - I_x\right) Q\right]\left(2V_T / b^2\bar{Q}S\right) \quad \text{(A-22)}$$

$$C_{NR} = \left[I_z\left(\frac{\partial \dot{R}}{\partial R}\right)\left(2V_T / b^2\bar{Q}S\right)\right. \quad \text{(A-23)}$$

Control Terms:

$$C_{X\delta i} = m\left(\frac{\partial \dot{U}}{\partial \delta_i}\right) / \bar{Q}S \quad \text{(A-24)}$$

$$C_{Y\delta i} = m\left(\frac{\partial \dot{V}}{\partial \delta_i}\right) / \bar{Q}S \quad \text{(A-25)}$$

$$C_{Z\delta i} = m\left(\frac{\partial \dot{W}}{\partial \delta_i}\right) / \bar{Q}S \quad \text{(A-26)}$$

$$C_{L\delta i} = I_x\left(\frac{\partial \dot{P}}{\partial \delta_i}\right) / \bar{Q}Sb \qquad \text{(A-27)}$$

$$C_{M\delta i} = I_y\left(\frac{\partial \dot{Q}}{\partial \delta_i}\right) / \bar{Q}S\bar{c} \qquad \text{(A-28)}$$

$$C_{N\delta i} = I_z\left(\frac{\partial \dot{R}}{\partial \delta_i}\right) / \bar{Q}Sb \qquad \text{(A-29)}$$

In the above, y and u are evaluated at their nominal values and the partials are obtained from the values of matrices in the linear model. This method ignores contributions to forces and moments due to $\dot{\alpha}$ and assumes that V_T is constant. Relaxation of these assumptions is possible, but the solution procedure is more difficult.

Two flight conditions were evaluated using this method. The constant coefficients (basic lift, drag, etc.) are determined by ensuring that $\dot{x} = 0$. Notice that the nondimensional coefficients at the two flight conditions are substantially different. This is most likely due to large control nonlinearities and the effect of flap deflection on basic aircraft characteristics and on horizontal tail nonlinearities.

FC1 is defined by:

V = 160 KIAS

h = 3500 feet

G = 0 (Gear Up)

Flaps = 15°

FC2 is defined by:

V = 140 KIAS

h = 1500 feet

G = 1 (Gear Down)

Flaps = 30°

Other assumptions

S = wing area = 1000 ft^2

b = wing span = 100 ft

$\bar{c}$ = average chord = 10 ft

m =

The resulting coefficients are shown below in Table A-1.

TABLE A-1. NONDIMENSIONAL COEFFICIENTS

FC1		FC2	
CXB		CXB	
0.11159E+01	ALPH	0.12661E+01	ALPH
0.33830E-01	Q	0.26276E-01	Q
0.11541E-01	THRUST_L	0.15070E-01	THRUST_L
0.11541E-01	THRUST_R	0.15070E-01	THRUST_R
0.71224E-03	D_STAB_L	0.52730E-03	D_STAB_L
0.71224E-03	D_STAB_R	0.52730E-03	D_STAB_R
0.00000E+00	D_RUDDER	0.00000E+00	D_RUDDER
0.33401E-03	D_ELEV_L	0.25056E-03	D_ELEV_L
0.33401E-03	D_ELEV_R	0.25056E-03	D_ELEV_R
0.38211E-03	D_AILE_L	0.28252E-03	D_AILE_L
0.38211E-03	D_AILE_R	0.28252E-03	D_AILE_R
-.61064E-03	D_SPLR_L	-.70234E-03	D_SPLR_L
-.61064E-03	D_SPLR_R	-.70234E-03	D_SPLR_R
-.91344E-01	K	-.16288E+00	K
CYB		CYB	
-.12764E+01	BETA	-.13282E+01	BETA
0.14274E+00	P	0.27300E+00	P
0.40748E+00	R	0.30631E+00	R
0.00000E+00	THRUST_L	0.00000E+00	THRUST_L
0.00000E+00	THRUST_R	0.00000E+00	THRUST_R
0.00000E+00	D_STAB_L	0.00000E+00	D_STAB_L
0.00000E+00	D_STAB_R	0.00000E+00	D_STAB_R
0.68085E-02	D_RUDDER	0.68867E-02	D_RUDDER
0.00000E+00	D_ELEV_L	0.00000E+00	D_ELEV_L
0.00000E+00	D_ELEV_R	0.00000E+00	D_ELEV_R
0.60524E-04	D_AILE_L	0.44354E-04	D_AILE_L
-.60524E-04	D_AILE_R	-.44354E-04	D_AILE_R
0.62347E-03	D_SPLR_L	0.85356E-03	D_SPLR_L
-.62347E-03	D_SPLR_R	-.85356E-03	D_SPLR_R
-.54122E-15	K	-.55180E-15	K
CZB		CZB	
-.69370E+01	ALPH	-.74416E+01	ALPH
-.38821E+00	Q	-.42188E+00	Q
-.16000E-04	THRUST_L	-.16445E-04	THRUST_L
-.16000E-04	THRUST_R	-.16445E-04	THRUST_R
-.83883E-02	D_STAB_L	-.84034E-02	D_STAB_L
-.83883E-02	D_STAB_R	-.84034E-02	D_STAB_R
-0.0000E+00	D_RUDDER	0.00000E+00	D_RUDDER
-.39325E-02	D_ELEV_L	-.39927E-02	D_ELEV_L
-.39325E-02	D_ELEV_R	-.39927E-02	D_ELEV_R
-.44963E-02	D_AILE_L	-.45015E-02	D_AILE_L
-.44963E-02	D_AILE_R	-.45015E-02	D_AILE_R
0.44794E-02	D_SPLR_L	0.60711E-02	D_SPLR_L
0.44794E-02	D_SPLR_R	0.60711E-02	D_SPLR_R
-.52182E+00	K	-.10091E+01	K

TABLE A-1. NONDIMENSIONAL COEFFICIENTS (Continued)

FC1		FC2	
CLB		CLB	
-.24436E+00	BETA	-.27875E+00	BETA
-.58055E+00	P	-.57419E+00	P
0.20669E+00	R	0.25012E+00	R
0.10871E-03	THRUST_L	0.14199E-03	THRUST_L
-.10871E-03	THRUST_R	-.14199E-03	THRUST_R
0.61458E-03	D_STAB_L	0.61540E-03	D_STAB_L
-.61458E-03	D_STAB_R	-.61540E-03	D_STAB_R
0.75713E-03	D_RUDDER	0.76630E-03	D_RUDDER
0.28741E-03	D_ELEV_L	0.29215E-03	D_ELEV_L
-.28741E-03	D_ELEV_R	-.29215E-03	D_ELEV-R
0.61045E-03	D_AILE_L	0.63462E-03	D_AILE_L
-.61045E-03	D_AILE_R	-.63462E-03	D_AILE_R
-.70755E-03	D_SPLR_L	-.96394E-03	D_SPLR_L
0.70755E-0E	D_SPLR_R	0.96394E-03	D_SPLR_R
-.10716E-15	K	-.10434E-15	K
CMB		CMB	
-.17373E+01	ALPH	-.24619E+01	ALPH
-.33986E+02	Q	-.33929E+02	Q
0.57743E-02	THRUST_L	0.75395E-02	THRUST_L
0.57743E-02	THRUST_R	0.75395E-02	THRUST_R
-.33534E-01	D_STAB_L	-.33518E-01	D_STAB_L
-.33534E-01	D_STAB_R	-.33518E-01	D_STAB_R
0.00000E+00	D_RUDDER	0.00000E+00	D_RUDDER
-.15602E-01	D_ELEV_L	-.15877E-01	D_ELEV_L
-.15602E-01	D_ELEV_R	-.15877E-01	D_ELEV_R
-.44041E-02	D_AILE_L	-.43994E-02	D_AILE_L
-.44041E-02	D_AILE_R	-.43994E-02	D_AILE_R
0.18332E-02	D_SPLR_L	0.23797E-02	D_SPLR_L
0.18332E-02	D_SPLR_R	0.23797E-02	D_SPLR_R
-.90779E-02	K	-.12474E+00	K
CNB		CNB	
0.13662E+00	BETA	0.14165E+00	BETA
-.13149E+00	P	-.13610E+00	P
-.14715E+00	R	-.15230E+00	R
0.18818E-02	THRUST_L	0.24579E-02	THRUST_L
-.18818E-02	THRUST_R	-.24579E-02	THRUST_R
0.13354E-03	D_STAB_L	0.12025E-03	D_STAB_L
-.13354E-03	D_STAB_R	-.12025E-03	D_STAB_R
-.26346E-02	D_RUDDER	-.26849E-02	D_RUDDER
0.76452E-04	D_ELEV_L	-.70973E-04	D_ELEV_L
-.76452E-04	D_ELEV_R	0.70973E-04	D_ELEV_R
0.14500E-03	D_AILE_L	0.15758E-03	D_AILE_L
-.14500E-03	D_AILE_R	-.15758E-03	D_AILE_R
-.35808E-03	D_SPLR_L	-.45348E-03	D_SPLR_L
0.35808E-03	D_SPLR_R	0.45348E-03	D_SPLR_R
0.17944E-15	K	0.20830E-15	K

APPENDIX B

DERIVATION OF OVERALL ERROR RATE EXPRESSIONS

This appendix derives expressions for overall system error probabilities for the aircraft path FDI system. These probabilities are based on the individual error probabilities of the hypothesis tests which are performed. First we define some fundamental events.

Let,

T_i = event of the i^{th} trigger indicating possible failure (B-1)

V_i = event of the i^{th} verify test choosing H_i over H_o (B-2)

$I_{i/j}$ = event of the i, j^{th} isolation test choosing H_i over H_j (B-3)

i,j, = 1, ..., 7 (i.e., seven control element failures)

Also, we define the aggregate events,

$$T = \bigcup_{i=1}^{7} Ti = \text{the event that some trigger indicates failure} \quad \text{(B-4)}$$

$$V = \bigcup_{i=1}^{7} v_i = \text{the event that some verify chooses Hi over } H_o \quad \text{(B-5)}$$

$$d_i = T \cap V_i \bigcap_{j \neq i} I_{i/j} = \text{the event that failure i is declared} \quad \text{(B-6)}$$

The relevant probabilities which we wish to calculate are,

P_{FA} = probability of choosing any failure mode i when there is no failure (B-7)

P_{FC} = probability of choosing H_i when H_j is true (B-8)

In order to compute the above, several intermediate probabilities are needed.

The probability of false trigger, P_{FT}, is the probability that T is true when there is no failure. Using Eq. B-4, we have

$$\overline{T} = \bigcap_{i=1}^{7} \overline{T}_i \tag{B-9}$$

and

$$P_{CR} = P(\overline{T}|Ho) = P(\overline{T}_i|H_o)^7 = [1-P(T_i|H_o)]^7 \tag{B-10}$$

assuming independent tests. Finally, $P_{FT} = 1-P_{CR}$.

Similarly, the probability of false verification, P_{FV}, is the probability that V is true under H_o and is computed from

$$P_{FV} = P(V|H_o = 1-[1-P(V_i|H_o)]^7 \tag{B-11}$$

again assuming independent tests.

In the following, the a priori probabilities of each failure mode are assumed to be equal, and all tests are assume to be independent.

The false alarm probability, P_{FA}, is given by

$$P_{FA} = P[\bigcup_{i\neq 0} d_i|H_o] \tag{B-12}$$

Since the events, d_i, are disjoint,

$$P_{FA} = \sum_i P(d_i|H_o)$$

$$= \sum_i P_{FT} \cdot P(V_i|H_o) \cdot \prod_{j=1} P(I_{i/j}|H_o) \qquad \text{(B-13)}$$

Assuming that $P(V_i|H_o)$ is the same for all i and $P(I_{i/j}|H_o)$ is the same for all j, we have

$$P_{FA} = 7\ P_{FT} \cdot P(V_i|H_o \cdot P(I_{i/j}|H_o)^6 \qquad \text{(B-14)}$$

Assuming that choosing H_i over H_j, $i \neq j \neq o$, when H_o is true is completely random, i.e., $P(I_{i/j}|H_o) = 1/2$, we have,

$$P_{FA} < \frac{7}{64} \cdot P_{FT} \cdot P(V_i|H_o) \qquad \text{(B-15)}$$

The probability of false classification, P_{FC}, is computed from $P_{FC} = 1-P_{CC}$ where

$$P_{CC} = P\left\{ \bigcup_{i\neq 0} d_i \cap H_i \;\middle|\; \bigcup_{i\neq 0} Hi \right\} \qquad \text{(B-16)}$$

Using Bayes rule and the fact that the events $d_i \cap H_i$ are disjoint, Eq. B-16 can be rewritten,

$$P_{CC} = \sum_j P(d_j|H_j) \cdot P(H_j| \bigcup_{j\neq 0} H_j) \qquad \text{(B-17)}$$

or when all of the a priori probabilities are equal,

$$P_{CC} = \frac{1}{7} \sum_{j=1}^{7} P(d_j|H_j) \qquad \text{(B-18)}$$

Now using Eq. B-6, the fact that each V_i and $I_{i/j}$ event is a subset of T, each test is independent, and the assumption that the individual test probabilities are equal for each failure mode, we have

$$P_{CC} = P(T|H_i)\ P(V_i|H_i) \cdot P(I_{i/j}|H_i)^6 \qquad \text{(B-19)}$$

$$\leq P(T_i|H_i)\ P(V_i|H_i)\ P(I_{i/j}|H_i)^6 \qquad \text{(B-20)}$$

Example

Suppose $P(T_i|H_i) = P(T_i|H_o) = P(V_i|H_o) = P(V_i|H_i) = P(I_{i/j}|H_j, j \neq i) = 10^{-4}$. Then we have,

$$P_{FT} < 10^{-3}$$

$$P_{FA} < 10^{-8}$$

$$P_{FC} < 10^{-3}$$

The false alarm spec, if expressed as 1 in 10^8 samples, represents better than 1 in 800 hours of flight time for a .03 second sample interval.

APPENDIX C

FLIGHT DATA PROCESSING DETAILS

The flight data values used by the Failure Detection and Isolation system are derived by converting the flight data of the tape into the required formats, by use of two external programs, CRUNCH and CONFIG. CRUNCH inputs the ASCII data and converts it into a binary, sequential file. CONFIG then converts the binary data into FDI required values and outputs them into another binary, sequential file. The conversions are as follows:

1. converts velocity from knots to feet/sec

2. computes the yaw rate by taking the mean of the two intermediate values from four sensor readings

3. the left and right thrust are computed from throttle positions based on a first-order system

   a. first, the previous thrust value is initialized

      ```
      prev_left_thrust  = .298 * left_throttle_pos
      prev_right_thrust = .298 * right_throttle_pos
      ```

   b. then, for every time step:

      ```
      average_throttle = (left_throttle_pos + right_throttle_pos)/2
      left_thrust = exp(-0.5 * delta_time)*prev_left-thrust + (exp(-0.5
      * delta_time)-1)*-2*.298*left_throttle

      prev_left_thrust = left_thrust
      right-thrust = exp(-0.5*delta_time)*prev_right_thrust + (exp(-0.5
      * delta_time)-1)*-2).298*right_throttle

      prev_right_thrust = right_thrust
      ```

4. invert normal acceleration measurement so that it is consistent with residual generation definition

5. compute stabilizer position from pilot units
   stab = 3 - pilot_unit

6. correct rudder for bias in measurement

   rudder = rudder - 5.6

7. make the aileron measurements complementary of each other

   left_aileron = - left_aileron
   right_aileron = right_aileron

8. compute spoiler positions

   left_spoiler = (left_spoil1 + left_spoil2)/2
   right_spoiler = (right_spoil1 + right_spoil2)/2

All other input values are channeled to the output file untouched. In addition, this program initializes the variable QBAR (dynamic pressure) to zero so that it can be computed in FDI. The output file is headed by an integer indicating the number of channels in the output file. This is followed by the channels themselves, listed by vectors of time. The first nine measured surface deflections are repeated at the end to serve as dummy commanded deflection values. A list of the output file channels is given below:

0 TIME
1 V_T
2 ALPHA
3 BETA
4 P
5 Q
6 R
7 AX
8 AY
9 AZ
10 QBAR
11 ALTITUDE

measured values:

12	LEFT THRUST
13	RIGHT THRUST
14	STABILIZER
15	STABILIZER
16	RUDDER
17	LEFT ELEVATOR
18	RIGHT ELEVATOR
19	LEFT AILERON
20	RIGHT AILERON
21	LEFT SPOILER
22	RIGHT SPOILER

commanded values:

23	LEFT THRUST
24	RIGHT THRUST
25	STABILIZER
26	STABILIZER
27	RUDDER
28	LEFT ELEVATOR
29	RIGHT ELEVATOR

REFERENCES

1. Looze, D.P., S. Krolewski, J.L. Weiss, N. Barrett, and J.S. Eterno, "Automatic Control Design Procedures for Restructurable Aircraft Control," NASA-CR-172489, NASA Langley Research Center Contract No. NAS1-17411, January 1985.

2. Weiss, J.L., and D.P. Looze, "Further Development of an Automatic Control Design Procedure for Restructurable Aircraft Control," TR-249, ALPHATECH, Inc., Burlington, MA.

3. Looze, D.L., et. al., "An Automatic Redesign Approach for Restructurable Control Systems," IEEE Control Systems Magazine, May 1985.

4. Weiss, J.L., and D.P. Looze, J.S. Eterno, D.B. Grunberg, "Initial Design and Evaluation of Automatic Restructurable Flight Control System Concepts," NASA-CR-178064 NASA Langley Research Center Contract #NAS1-17411, June 1986.

5. Weiss, J.L., A.S. Willsky, D.P. Looze, J.T. Crawford, and R.R. Huber, "Detection and Isolation of Control Surface Effectiveness Failures in High Performance Aircraft," in IEEE Proceedings of NAECON, May 19853.

6. Huber, R., and B. McCullough, "Self-Repairing Flight Control System, Society of Automotive Engineers, Paper #SAE-841552, presented at Aerospace Congress and Exposition, Long Beach, CA., Oct. 1984.

7. Weiss, J.L., J.M. Stifel, and K.S. Govindaraj, "Flight Test Results of a Control Element FDI Algorithm," Presented at NAECON '86, Dayton, Ohio.

8. Weiss, J.L., J.S. Eterno, and J.Y. Hsu, "Evaluation of Detectability and Distinguishability of Aircraft Control Element Failures Using Flight Test Data," 1986 American Control Conference, NASA Contract NAS1-18004, TP-235, ALPHATECH, Inc., Burlington, Massachusetts, March 1986.

9. Bonice, W. et al., AIAA Paper No. 85-1973, Conference of Guidance, Navigation, and Control, August 1985, Snowmass, Colorado.

10. Bonnice, W., et al., "The Evaluation of the OSGLR Algorithm for Restructurable Control," NASA-CR-178083, NASA Contract NAS1-17556, May 1986.

11. Montoya, R.J. et el., NASA Conference Publication 227, NASA Langley Research Center, Hampton, Virginia, September 21-22, 1982.

REFERENCES (Continued)

12. Ropelewski, R., Aviation Week and Space Technology, August 5, 1985.

13. Lyon, M., Airline Executive, August 1985.

14. Deckert, J.C., M.N. Desai, J.J. Deyst, and A.S. Willsky, "F8-DFBW Failure Identification Using Analytic Redundancy," IEEE Trans. Automatic Control, Vol. AC-22, No. 5, October 1977, pp. 795-803.

15. Deckert, J.C., "Flight Tests Results for the F-8 Digital Fly-by-Wire Aircraft Control Sensor Analytic Redundancy Management Technique," Proc. AIAA Guidance and Control Conference, Albuquerque, New Mexico, August 1981, pp. 491-499.

16. Deckert, J.C., Desai, M.N., Deyst, J.J., and Willsky, A.S., "Reliable Dual-Redundant Sensor Failure Detection and Identification for the NASA F-8 DFBW Aircraft," NASA CR-2944, February 1978.

17. Willsky, A.S., "A Survey of Design Methods for Failure Detection in Dynamic Systems," Automatica, Vol. 12, 1976, pp. 601-611; also in AGARDograph No. 224 on "Integrity in Electronic Flight Control Systems."

18. Chow, E.Y., and A.S. Willsky, "Issues in the Development of a General Design Algorithm for Reliable Failure Detection," Proc. of IEEE Conference on Decision and Control, Albuquerque, New Mexico, December 1980.

19. Lou, X.-C, A.S. Willsky, and G.C. Verghese, "Failure Detection with Uncertain Models," Preprint Sept. 1982, invited paper for 1983 American Control Conference, June 1983.

20. Chow, E.Y., and A.S. Willsky, "Analytical Redundancy and the Design of Robust Failure Detection System," IEEE Transactions on Automatic Control, July 1984.

21. Beard, R.V., "Failure Accommodation in Linear Systems Through Self-Reorganization," Report MVT-71-1, Man Vehicle Lab, MIT, Cambridge, Massachusetts, February 1971.

22. Pattipati, K.R., A.S. Willsky, J.C. Deckert, J.S. Eterno, and J.L. Weiss, "A Design Methodology for Robust Failure Detection and Isolation," 1984 American Control Conference, San Diego, CA, June 1984.

23. Weiss, J.L., A.S. Willsky, K.R. Pattipati, and J.S. Eterno, "Application of FDI Metrics to Detection and Isolation of Sensor Failures in Turbine Engines, " 1984 American Control Conference, San Diego, CA, June 1984.

24. Weiss, J.L., K.R. Pattipati, A.S. Willsky, J.S. Eterno, and J.T. Crawford, "Robust Detection/Isolation/Accommodation for Sensor Failures," NASA-CR-174797, NASA Lewis Research Center Contract No. NAS3-24078, September 1984.

REFERENCES (Continued)

25. Willsky, A.S., "Failure Detection in Dynamic Systems," AGARD Lecture Series No. 109 on "Fault Tolerance Design and Redundancy Management Techniques," Athens, Rome, and London, Oct. 1980.

26. Jones, H.L., "Failure Detection in Linear Systems," PhD Thesis, Dept. of Aeronautics and Astronautics, MIT, Cambridge, Massachusetts, September 1973.

27. Messerole, J.S., Jr., "Detection Filters for Fault-Tolerant Control of Turbofan Engines," PhD Thesis, MIT, Dept. of Aeronautics and Astronautics, June 1981.

28. Gilmore, J.P., and R.A. McKern, "A Redundant Strapdown Inertial Reference Unit (SIRU)," J. Spacecr. Rockets, January 1972.

29. Etkin, B., Dynamics of Atmospheric Flight, Wiley & Sons, Inc., NY 1972.

30. Lou, X.-C., "A System Failure Detection Method: The Failure Projection Method," S.M. Thesis, MIT Dept. of Elec. Eng. and Comp. Sci., Rept. LIDS-TH-1203, MIT Lab for Inf. and Dec. Sys., June 1982.

31. Vander Velde, W., and A.M. San MArtin, "Design of Robust Failure Detection Filters," Proceedings of 1986 Automatic Control Conference, Seattle, Washington, June 18-20, 1986.

32. Van Trees, H.L., Detection Estimation and Modulation Theory, Part I, Wiley & Sons, Inc., NY, 1968.

33. Wald, A., Sequential Analysis, Dover Publications, Inc., NY, 1947.

34. Craig, A.T., and Hogg, R.V., Introduction to Mathematical Statistics, Macmillan Publishing Co., NY, 1970.

35. Tantaratana, S.," Comparison of the SPRT and the Sequential Linear Detector in Autoregressive Noise," IEEE Transactions on Information Theory, Vol, IT-31, Sept. 1985.

36. Weiss, Jerold L., J.Y. Hsu, and Willsky, Alan S., "Design and Evaluation of a Failure Detection and Isolation Algorithm for Restructurable Aircraft Control," Technical Progress Report for NAS1-18004.

37. Doyle, J.C., and G. Stein, "Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis," IEEE Transactions on Automatic Control, Vol. AC-16, December 1971, pp. 529-552.

38. Naeini, A.E., et al., "Robust Detection, Isolation, and Accommodation for Sensor Failures," Proceedings of 1985 American Control Conference, Boston, Massachusetts.

REFERENCES (Continued)

39. Tou, J.T., and R.C. Gonzalez, Pattern Recognition Principles, Addison Wesley Publishers, Reading, MA, 1974.

40. Fukunaga, K., Introduction to Statistical Pattern Recognition, Academic Press, NY, 1972.

41. Hawkes, R.M., and J.B. Moore, "Performance Bounds for Adaptive Estimation," Proc. IEEE, Vol. 64, No. 8, pp. 1143-1150, August 1976.

42. Anderson, B.D.O., J.B. Moore, and R.M. Hawkes, "Model Approximations via Prediction Error Identification," Automatica, Vol. 14, No. 6, pp. 615-622, November 1978.

43. Baram, Y., and N. Sandell, Jr., "Consistent Estimation of Finite Parameter Sets with Application to Linear System Identification," IEEE Transactions on Automatic Control, Vol. AC-23, pp. 451-454, June 1978.

44. Kailath, T., "The Divergence and Bhattacharyya Distance Measures in Signal Selection," IEEE Transactions on Communication Technology, Vol. COM-15, No. 1, pp. 52-60, February 1967.

45. Kazakos, D., and P. Papantoni-Kazakos, "Spectral Distance Measures Between Gaussian Processes," IEEE Transactions on Automatic Control, Vol. 25, No. 5, pp. 950-959, October 1979.

46. Kazakos, D., "On Resolution and Exponential Discrimination Between Gaussian Stationary Vector Processes and Dynamic Models," IEEE Transactions on Automatic Control, Vol. 25, No. 2, pp. 294-296, April 1980.

47. Bhattacharyya, A., "On a Measure of Divergence Between Two Statistical Populations Defined by Their Probability Distributions," Bull. Calcutta, Math. Soc., Vol. 35, pp. 99-109, 1943.

48. Kullback, S., Information Theory and Statistics, Dover Pub., NY, 1959.

49. Kazakos, D., et al., IEEE Transactions on Automatic Control, Vol. 25, No. 5, pp. 950-959, October 1979.

50. Deckert, J.C., and J. Deyst, "Maximum Likelihood Failure Detection Techniques Applied to the Shuttle RCS Jets," Journal of Spacecraft and Rockets, Vol. 13, No. 2, February 1976.

51. Caglayan, A., et al., "Reconfiguration in Stages, Part I," Proceedings of NAECON, 19-23 May 1986.

52. Motyka, P., et al., "A Comparison of Aircraft Control Sun Force Failure Detection and Isolation Algorithms," Proceedings of ACC, 18-20 June 1986.

REFERENCES (Continued)

53. Gross, H., et al., "Renewed Interest in Hinge Moment Models for Failure Detection and Isolation," Proceedings of 1986 ACC, June 1986.

54. Kay, S.M. et al., Proceedings of the IEEE, Vol. 69, No. 11, November 1981.

55. Oppenheim, Alan V., and R.W. Shafer, Digital Signal Processing, Prentice Hall, Englewood Cliffs, NJ, 1975.

56. Weiss, J.L., "Integrated Restructurable Flight Control System Development, Second Quarterly Report, Demonstration System Definition," NASA Contract No. NAS1-17411, TR-302, ALPHATECH, Inc., Burlington, Massachusetts, August 1986.

57. Davenport, WB, Probability and Random Processes, McGraw Hill, New York, New York, 1970.

58. Weiss, J.L., D.P. Looze, J.S. Eterno, and D.B. Grunberg, "Initial Design and Evaluation of Automatic Restructurable Flight Control System Concepts, NASA Contractor Report 178064 under NASA Contract No. NAS1-17411, TR-269-1, ALPHATECH, Inc., Burlington, Massachusetts, June 1986.

59. Kay, S.M., and S.L. Marple, Jr., "Spectrum Analysis — A Modern Perspective," Proc. IEEE, Vol. 69, No. 11, pp. 1380-1419, November, 1981.

Standard Bibliographic Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
NASA CR 178213		

4. Title and Subtitle	5. Report Date
Design and Evaluation of a Failure Detection and Isolation Algorithm for Restructurable Control Systems	March 1987
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Jerold L. Weiss, John Y. Hsu	TR-309
9. Performing Organization Name and Address	10. Work Unit No.
ALPHATECH, Inc. 2 Burlington Executive Center 111 Middlesex Turnpike Burlington, MA 01803	11. Contract or Grant No. NAS1-18004
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered
NASA Langley Research Center Hampton, Virginia 23665	Contractor Report
	14. Sponsoring Agency Code

15. Supplementary Notes

Annual Report

16. Abstract

This effort has developed and explored the use of a decentralized approach to failure detection and isolation for use in restructurable control systems. This work has produced,

1. A method for evaluating fundamental limits to FDI performance,
2. Application of this method using flight recorded data,
3. A working control element FDI system with maximal sensitivity to critical control element failures,
4. Extensive testing of this system on realistic simulations, and
5. A detailed design methodology for this system which involves parameter optimization (with respect to model uncertainties) and sensitivity analyses.

For this project, we have concentrated on detection and isolation of generic control element failures since these failures frequently lead to emergency conditions and since knowledge of remaining control authority is essential for control system redesign. The failures we considered are generic in the sense that no temporal failure signature information was assumed. Thus, various forms of "functional failures" are treated in a unified fashion. Such a treatment results in a robust FDI system (i.e., one that covers all failure modes) but sacrifices some performance when detailed failure signature information is known, useful, and employed properly. We assumed throughout this project that all sensors are validated (i.e., contain only in-spec errors) and that only the first failure of a single control element needs to be detected and isolated. The FDI system which has been developed will handle a class of multiple failures.

17. Key Words (Suggested by Authors(s))	18. Distribution Statement
FDI, Reconfigurable Control, Analytic Redundancy, Failure Detection	

19. Security Classif.(of this report)	20. Security Classif.(of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	253	

For sale by the National Technical Information Service, Springfield, Virginia 22161